

PCT/JP97/03239

日本国特許庁 12.09.97
 PATENT OFFICE
 JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
 Date of Application:

1996年 9月13日

REC'D 31 OCT 1997

PCT

出願番号
 Application Number:

平成 8年特許願第243060号

出願人
 Applicant(s):

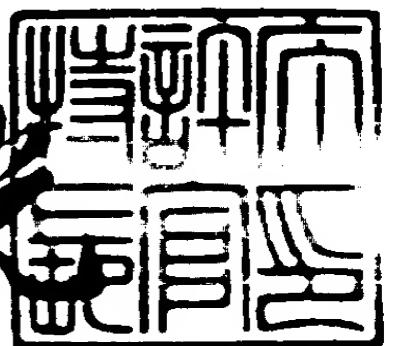
財団法人相模中央化学研究所
 株式会社プロテジーン

PRIORITY DOCUMENT

1997年10月17日

Patent Office

荒井寿光



【書類名】 特許願
【整理番号】 S018065
【提出日】 平成 8年 9月13日
【あて先】 特許庁長官殿
【発明の名称】 分泌シグナル配列を有するヒト蛋白質およびそれをコードするDNA
【請求項の数】 4
【発明者】
【住所又は居所】 神奈川県相模原市南台1-9-2
【氏名】 加藤 誠志
【発明者】
【住所又は居所】 神奈川県相模原市西大沼4-4-1
【氏名】 関根 伸吾
【発明者】
【住所又は居所】 東京都葛飾区高砂5-13-11
【氏名】 山口 知子
【発明者】
【住所又は居所】 神奈川県藤沢市長後647-2
【氏名】 小林 みどり
【特許出願人】
【代表出願人】
【識別番号】 000173762
【郵便番号】 229
【住所又は居所】 神奈川県相模原市西大沼4丁目4番1号
【氏名又は名称】 財団法人相模中央化学研究所
【代表者】 近藤 聖
【電話番号】 0427(42)4791
【特許出願人】
【郵便番号】 153

【住所又は居所】 東京都目黒区中町2丁目20番3号

【氏名又は名称】 株式会社プロテジーン

【代表者】 棚井 丈雄

【電話番号】 03(3792)1019

【手数料の表示】

【予納台帳番号】 011501

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 分泌シグナル配列を有するヒト蛋白質およびそれをコードするDNA

【特許請求の範囲】

【請求項1】 配列番号1から配列番号9で表されるアミノ酸配列のいずれかを含む蛋白質。

【請求項2】 請求項1記載の蛋白質のいずれかをコードするDNA。

【請求項3】 配列番号10から配列番号18で表される塩基配列のいずれかを含むcDNA。

【請求項4】 配列番号19から配列番号27で表される塩基配列のいずれかからなる、請求項3記載のcDNA。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、分泌シグナル配列を有するヒト蛋白質、およびそれをコードしているDNAに関する。本発明の蛋白質は、医薬品として、あるいは該蛋白質に対する抗体を作製するための抗原として用いることができる。本発明のヒトcDNAは、遺伝子診断用プローブや遺伝子治療用遺伝子源として用いることができる。また、該cDNAがコードしている蛋白質を大量生産するための遺伝子源として用いることができる。

【0002】

【従来技術】

細胞は多くの蛋白質を細胞外に分泌している。これらの分泌蛋白質は、細胞の増殖制御、分化誘導、物質輸送、生体防御などにおいて重要な役割を演じている。分泌蛋白質は細胞内蛋白質と異なり、細胞外で作用するので、注射や点滴などによる体内投与が可能であり、医薬としての可能性を秘めている。事実、インターフェロン、インターロイキン、エリスロポイエチン、血栓溶解剤など、多くのヒト分泌蛋白質が現在医薬として使用されている。また、これら以外の分泌蛋白質についても臨床試験が進行中であり、医薬品を目指した用途開発がなされている。

る。ヒト細胞は、まだ多くの未知の分泌蛋白質を生産していると考えられており、これらの分泌蛋白質並びにそれをコードしている遺伝子が入手できれば、これらを用いた新しい医薬品開発が期待できる。

【0003】

従来、これらの分泌蛋白質を得るために、大量の血液や細胞培養上澄などから生理活性を指標としてターゲット蛋白質を単離精製し、その一次構造を決定したのち、得られたアミノ酸配列情報に基づいて対応するcDNAをクローニングし、これを用いて組換え蛋白質を生産する方法がとられてきた。しかし、分泌蛋白質は一般に含有量が低いので、単離精製することが困難なものが多い。一方、分泌蛋白質やI型膜蛋白質は、そのアミノ末端(N末端)に分泌シグナル配列と呼ばれる約20アミノ酸残基からなる疎水性の配列を有している。したがって、この分泌シグナル配列の有無を指標に、分泌蛋白質やI型膜蛋白質をコードする遺伝子をクローニングすることができると考えられる。

【0004】

【発明が解決しようとする課題】

本発明の目的は、分泌シグナル配列を有する新規なヒト蛋白質、および該蛋白質をコードするDNAを提供することである。

【0005】

【課題を解決するための手段】

本発明者らは銳意研究の結果、ヒト完全長cDNAバンクの中から分泌シグナル配列を有するcDNAを選択的にクローニングし、本発明を完成した。すなわち、本発明は分泌シグナル配列を有するヒト蛋白質である、配列番号1から配列番号9で表されるアミノ酸配列のいずれかを含む蛋白質を提供する。また本発明は上記蛋白質をコードするDNA、例えば配列番号10から配列番号18で表される塩基配列のいずれかを含むcDNAを提供する。

【0006】

本発明の蛋白質は、ヒトの臓器、細胞株などから単離する方法、本発明のアミノ酸配列に基づき化学合成によってペプチドを調製する方法、あるいは本発明の

ヒト分泌蛋白質をコードするDNAを用いて組換えDNA技術で生産する方法などにより取得することができるが、組換えDNA技術で取得する方法が好ましく用いられる。例えば、本発明のcDNAを有するベクターからインビトロ転写によってRNAを調製し、これを鑄型としてインビトロ翻訳を行なうことによりインビトロで発現できる。また翻訳領域を公知の方法により適当な発現ベクターに組換えてやれば、大腸菌、枯草菌、酵母、動物細胞等で、コードしている蛋白質を大量に発現させることができる。

【0007】

本発明の蛋白質を、大腸菌などの微生物で発現させる場合には、微生物中で複製可能なオリジン、プロモーター、リボソーム結合部位、cDNAクローニング部位、ターミネーター等を有する発現ベクターに、本発明のcDNAの翻訳領域を組換えた発現ベクターを作成し、該発現ベクターで宿主細胞を形質転換したのち、得られた形質転換体を培養してやれば、該cDNAがコードしている蛋白質を微生物内で大量生産することができる。この際、分泌シグナル配列を除去した翻訳領域に開始コドンを付加して発現させてやれば、成熟蛋白質を得ることができる。あるいは、他の蛋白質との融合蛋白質として発現させることもできる。該融合蛋白質を適当なプロテアーゼで切断することによって該cDNAがコードする蛋白質部分のみを取得することもできる。

【0008】

本発明の蛋白質を、動物細胞で分泌発現させる場合には、該cDNAの翻訳領域を、プロモーター、スプライシング領域、ポリ(A)付加部位等を有する動物細胞用発現ベクターに組換え、動物細胞内に導入してやれば、本発明の蛋白質を成熟蛋白質として細胞外に分泌生産することができる。

【0009】

本発明の蛋白質には、配列番号1から配列番号9で表されるアミノ酸配列のいかなる部分アミノ酸配列を含むペプチド断片(5アミノ酸残基以上)も含まれる。これらのペプチド断片は抗体を作製するための抗原として用いることができる。また、本発明の蛋白質は、シグナル配列が除去された後、成熟蛋白質の形で細胞外に分泌される。したがって、これらの成熟蛋白質は本発明の蛋白質の範疇に

はいる。成熟蛋白質のN末端アミノ酸配列は、シグナル配列切断部位決定法〔特開平8-187100〕を用いて容易に求めることができる。また、多くの分泌蛋白質は、分泌された後プロセシングを受けて活性型となる。このような活性型となった蛋白質あるいはペプチドも本発明の蛋白質の範疇にはいる。アミノ酸配列の中に糖鎖結合部位が存在すると、適当な動物細胞で発現させれば糖鎖が付加した蛋白質が得られる。したがって、このような糖鎖が付加した蛋白質あるいはペプチドも本発明の蛋白質の範疇にはいる。

【0010】

本発明のDNAには、上記蛋白質をコードするすべてのDNAが含まれる。該DNAは、化学合成による方法、cDNAクローニングによる方法などを用いて取得することができる。

【0011】

本発明のcDNAは、例えばヒト細胞由来cDNAライブラリーからクローニングすることができる。cDNAはヒト細胞から抽出したポリ(A)⁺RNAを鑄型として合成する。ヒト細胞としては、人体から手術などによって摘出されたものでも培養細胞でも良い。cDNAは、岡山-Berg法〔Okayama, H. and Berg, P., Mol. Cell. Biol. 2: 161-170 (1982)〕、Gubler-Hoffmann法〔Gubler, U. and Hoffmann, J. Gene 25: 263-269 (1983)〕などいかなる方法を用いて合成してもよいが、完全長クローニングを効率的に得るためにには、実施例にあげたようなキャッピング法〔Kato, S. et al., Gene 163: 193-196 (1995)〕を用いることが望ましい。

【0012】

分泌シグナル配列を有するヒト蛋白質をコードするcDNAの第一次選別は、DNAライブラリーから任意に選択したcDNAクローニングの部分塩基配列決定内の疎水性部分の有無の確認によって行なう。次いでシーケンシングによる全塩基配列の決定、インビトロ翻訳による蛋白質発現によって第二次選別行なう。本

発明のcDNAが、分泌シグナル配列を有する蛋白質をコードしていることの確認は、シグナル配列検出法 [Yokoyama-Kobayashi, M. et al.、Gene 163: 193-196 (1995)] を用いて行う。すなわち、ターゲット蛋白質のN末端をコードするcDNA断片を、ウロキナーゼのプロテアーゼドメインをコードするcDNAと融合させたのち、COS細胞内で発現させ、ウロキナーゼ活性が細胞培養液中に検出された場合には、挿入したcDNA断片がコードしている部分が、シグナル配列として機能していることを意味する。

【0013】

本発明のcDNAは、配列番号10から配列番号18で表される塩基配列あるいは配列番号19から配列番号27で表される塩基配列のいずれかを含むことを特徴とするものである。それぞれのクローン番号(HP番号)、cDNAクローンが得られた細胞、cDNAの全塩基数、コードしている蛋白質のアミノ酸残基数をそれぞれ表1にまとめて示した。

【0014】

【表1】

表1

配列番号	HP番号	細胞	塩基数	アミノ酸 残基数
1、10、19	HP00658	HT-1080	1296	154
2、11、20	HP00714	KB	3311	315
3、12、21	HP00876	胃癌	1152	158
4、13、22	HP01134	肝臓	1749	376
5、14、23	HP10029	KB	988	173
6、15、24	HP10189	KB	390	93
7、16、25	HP10269	U937	4667	1172
8、17、26	HP10298	胃癌	1086	122

9、18、27 HP10368 胃癌

866

175

【0015】

なお、配列番号19から配列番号27のいずれかに記載のcDNAの塩基配列に基づいて合成したオリゴヌクレオチドプローブを用いて、本発明で用いたヒト細胞株やヒト組織から作製したcDNAライブラリーをスクリーニングすることにより、本発明のcDNAと同一のクローンを容易に得ることができる。

【0016】

一般にヒト遺伝子は個体差による多型が頻繁に認められる。従って配列番号10から配列番号27において、1又は複数個のヌクレオチドの付加、欠失および／又は他のヌクレオチドによる置換がなされているcDNAも本発明の範疇にはいる。

【0017】

同様に、これらの変更によって生じる、1又は複数個のアミノ酸の付加、欠失および／又は他のアミノ酸による置換がなされている蛋白質も、配列番号1から配列番号9で表されるアミノ酸配列を有するそれぞれの蛋白質の活性を有する限り、本発明の範疇に入る。

【0018】

本発明のcDNAには、配列番号10から配列番号18で表される塩基配列あるいは配列番号19から配列番号27で表される塩基配列のいかなる部分塩基配列を含むcDNA断片(10bp以上)も含まれる。例えば、実施例に示した様に、分泌シグナル配列をコードする部分は、他の蛋白質をコードするcDNAと融合させることにより、任意の蛋白質を細胞外に分泌させるための手段として用いることができる。また、センス鎖およびアンチセンス鎖からなるDNA断片もこの範疇にはいる。これらのDNA断片は遺伝子診断用のプローブとして用いる可以である。

【実施例】

次に実施例により発明を具体的に説明するが、本発明はこれらの例に限定され

るものではない。DNAの組換えに関する基本的な操作および酵素反応は、文献[“Molecular Cloning. A Laboratory Manual”, Cold Spring Harbor Laboratory, 1989]に従った。制限酵素および各種修飾酵素は特に記載の無い場合宝酒造社製のものを用いた。各酵素反応の緩衝液組成、並びに反応条件は付属の説明書に従った。cDNA合成は文献[Kato, S. et al., Gene 150: 243-250 (1994)]に従った。

【0020】

(1) ポリ(A)⁺RNAの調製

mRNAを抽出するためのヒト細胞として、纖維肉腫細胞株HT-1080(ATCC CCL 121)、類表皮癌細胞株KB(ATCC CRL 17)、ホルボールエステルで刺激した組織球リンホーマ細胞株U937(ATCC CRL 1593)、手術によって摘出された胃癌組織並びに肝臓を用いた。それぞれの細胞株の培養は、常法に従って行った。

【0021】

ヒト細胞約1gを5.5Mグアニジウムチオシアネート溶液20ml中でホモジナイズした後、文献[Okayama, H. et al., “Methods in Enzymology” Vol. 164, Academic Press, 1987]に従い、総mRNAを調製した。これを20mMトリス塩酸緩衝液(pH7.6)、0.5M NaCl、1mM EDTAで洗浄したオリゴdTセルロースカラムにかけ、上掲文献に従いポリ(A)⁺RNAを得た。

【0022】

(2) cDNAライブラリーの作製

上記ポリ(A)⁺RNA 10 μgを100mMトリス塩酸緩衝液(pH8)に溶解し、RNaseを含まないバクテリア由来アルカリホスファターゼ1単位を添加し、37°C 1時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを50mM 酢酸ナトリウム(pH6)、1mM EDTA、0.1%2-メルカプトエタノール、0.01%Triton X-100溶液に溶解した。これに、タバコ由来酸ピロホスファターゼ(エピセンターテクノ

ロジーズ社製) 1 単位を添加して、総量 $100 \mu l$ で $37^\circ C$ 1 時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを水に溶解し、脱キップ処理したポリ (A) ⁺RNA 溶液を得た。

【0023】

脱キップ処理したポリ (A) ⁺RNA、DNA-RNA キメラオリゴヌクレオチド (5' - dG - dG - dG - dG - dA - dA - dT - dT - dC - dG - dA - G - G - A - 3') 3 nmol を 50 mM トリス塩酸緩衝液 (pH 7.5)、0.5 mM ATP、5 mM MgCl₂、10 mM 2-メルカプトエタノール、25% ポリエチレンリコール水溶液に溶解し、T4 RNA リガーゼ 50 単位を添加し、総量 $30 \mu l$ で $20^\circ C$ 12 時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを水に溶解し、キメラオリゴキップ付加ポリ (A) ⁺RNA を得た。

【0024】

本発明者らが開発したベクター pKA1 (特開平4-117292号公報) を KpnI で消化後、末端転移酵素により約 60 個の dT テールを付加した。これを EcoRV 消化して片側の dT テールを除去したものをベクタープライマーとして用いた。

【0025】

先に調製したキメラオリゴキップ付加ポリ (A) ⁺RNA 6 μg を、ベクタープライマー 1.2 μg とアニールさせた後、50 mM トリス塩酸緩衝液 (pH 8.3)、75 mM KC1、3 mM MgCl₂、10 mM ジチオスレイトール、1.25 mM dNTP (dATP + dCTP + dGTP + dTTP) 溶液に溶解し、逆転写酵素 (GIBCO-BRL 社製) 200 単位を添加し、総量 $20 \mu l$ で $42^\circ C$ 1 時間反応させた。反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを 50 mM トリス塩酸緩衝液 (pH 7.5)、100 mM NaCl、10 mM MgCl₂、1 mM ジチオスレイトール溶液に溶解した。

反応液をフェノール抽出後、エタノール沈殿を行ない、ペレットを 20 mM トリス塩酸緩衝液 (pH 7.5)、100 mM KC1、4 mM MgCl₂、10

$5 \text{mM} (\text{NH}_4)_2\text{SO}_4$ 、 $50 \mu\text{g}/\text{ml}$ 牛血清アルブミン溶液に溶解した。これに大腸菌DNAリガーゼ60単位を添加し、 16°C 16時間反応させた。反応液に $2 \text{mM dNTP} 2 \mu\text{l}$ 、大腸菌DNAポリメラーゼI 4単位、大腸菌RNase H₀ 1単位を添加し、 12°C 1時間ついで 22°C 1時間反応させた。

【0026】

次いでcDNA合成反応液を用いて大腸菌DH12S (GIBCO-BRL社製) の形質転換を行なった。形質転換はエレクトロポレーション法によって行なった。形質転換体の一部を $100 \mu\text{g}/\text{ml}$ アンピシリン含有 $2 \times \text{YT}$ 寒天培地上に蒔いて 37°C 一晩培養した。寒天上に生じた任意のコロニーを拾い $100 \mu\text{g}/\text{ml}$ アンピシリン含有 $2 \times \text{YT}$ 培地 2ml に接種して 37°C で一晩培養した。培養液を遠心して、菌体からアルカリリシス法によりプラスミドDNAを調製した。プラスミドDNAはEcoRIとNotIで二重消化した後、0.8%アガロースゲル電気泳動を行ないcDNAインサートの大きさを求めた。また、得られたプラスミドを鋳型にして、蛍光色素で標識したM13ユニバーサルプライマーとTaqポリメラーゼ (アプライドバイオシステムズ社製キット) を用いてシーケンス反応を行なった後、蛍光DNAシーケンサー (アプライドバイオシステムズ社) にかけてcDNAの5'末端約400bpの塩基配列を決定した。配列データはホモ・プロテインcDNAバンクデータベースとしてファイル化した。

【0027】

(3) 分泌シグナル配列を有する蛋白質をコードしているcDNAの選択

ホモ・プロテインcDNAバンクに登録された塩基配列を3フレームのアミノ酸配列に変換し、開始コドンから始まるオープンリーディングフレーム (ORF) の有無を調べた。次いでORFがコードしている部分のN末端に分泌蛋白質に特有なシグナル配列が認められるものを選択した。これらのクローンについては、エキソヌクレアーゼIIIによる欠失法を用いて、5'並びに3'両方向からシーケンシングを行い、全塩基配列の決定を行った。ORFがコードしている蛋白質について、Kyte-Doolittleの方法 [Kyte, J. & Doolittle, R. F., J. Mol. Biol. 157: 105-132

(1982)]により、疎水性／親水性プロフィールを求め、疎水性領域の有無を調べた。コードしている蛋白質のアミノ酸配列中に膜貫通ドメインと思われる疎水的な領域がない場合には、この蛋白質は分泌蛋白質あるいは膜貫通ドメインを持たない膜蛋白質であると見なした。

【0028】

(4) 分泌シグナル検出ベクター pSSD3 の構築

SV40プロモーターとウロキナーゼのプロテアーゼドメイン cDNA を有する pSSD1 [Yokoyama-Kobayashi, M. et al., Gene 163: 193-196 (1995)] 1 μ gを、BglII 5単位とEcoRV 5単位で消化した。ついでCIP処理によって5'末端の脱リン酸化を行った後、アガロースゲル電気泳動によって約4.2 kbp のDNA断片をゲルから切り出し精製した。

【0029】

2本のオリゴDNAリンカー L1 (5' - GATCCGGGTCACGTTGGAT - 3') と L2 (5' - ATCCCACGTTGACCCGG - 3') を合成し、T4ポリヌクレオチドキナーゼによりリン酸化した。両者をアニールしたのち、先に調製した pSSD1 の切断片と T4 DNA リガーゼにより連結し、大腸菌JM109を形質転換した。形質転換体からプラスミド pSSD3 を調製し、リンカー挿入部分の塩基配列を決定することにより目的とする組換え体を確認した。得られたプラスミドの構造を図1に示す。本プラスミドベクターは、ポリクローニング部位に3種の平滑末端生成制限酵素部位、SmaI、PmaCI、EcoRV を有している。これらの切断部位は 7 bp の間隔で並んでいるので、この中のいずれかを選べば、挿入する cDNA 断片の3種のフレームと合わせて融合蛋白質を発現するベクターを構築できる。

【0030】

(ニ) 分泌シグナル配列の機能確認

坂が分泌シグナル配列として機能することを、文献記載の方法 [Yokoyama-Kobayashi, M. et al., Gene 163: 193-1

96(1995)]によって確認した。まずターゲットcDNAを含んでいるプラスミドを、分泌シグナル配列をコードしていると考えられる部分の下流に存在する適当な制限酵素部位で切断した。もしこの制限酵素部位が5'突出末端である場合には、クレノウ処理によって平滑末端にした。さらにHindIIIによる消化を行い、SV40プロモーターとその下流に分泌シグナル配列をコードしているcDNAを含むDNA断片をアガロースゲル電気泳動によって単離した。この断片を、pSSD3のHindIIIと、ウロキナーゼのコーディングフレームと合うように選択した制限酵素部位の間に挿入し、ターゲットcDNAの分泌シグナル配列部分とウロキナーゼプロテアーゼドメインの融合蛋白質を発現するためのベクターを構築した(図2参照)。

【0031】

融合蛋白質発現ベクターを有する大腸菌(宿主: JM109)を100 μ g/mlアンピシリン含有2×YT培地2ml中で37℃2時間培養した後、ヘルペスファージM13KO7(50 μ l)を添加し、37℃で一晩培養した。遠心によって分離した上澄からポリエチレンリコール沈殿によって一本鎖ファージ粒子を得た。これを100 μ lの1mMトリス-0.1mMEDTA、pH8(TE)に懸濁した。また対照として、pSSD3、並びにウロキナーゼの完全長cDNAを含むベクターpKA1-UPA[Yokoyama-Kobayashi, M. et al., Gene 163: 193-196(1995)]から同様にして調製した一本鎖ファージ粒子懸濁液を用いた。

【0032】

サル腎臓由来培養細胞COS7は、10%ウシ胎児血清を含むダルベッコ改変イーグル(DMEM)培地中、5%CO₂存在下、37℃で培養した。1×10⁵個のCOS7細胞を6穴プレート(ヌンク社、穴の直径3cm)に植え、5%CO₂存在下、37℃で22時間培養した。培地除去後、リン酸緩衝液で細胞表面を洗浄し、さらに50mMトリス塩酸(pH7.5)を含むDMEM(TDEM)で再度洗浄した。この細胞に一本鎖ファージ懸濁液1 μ l、DMEM培地0.6ml、TRANSFECTAMTM(IBF社)3 μ lを懸濁したものを添加し、5%CO₂存在下、37℃で3時間培養した。サンプル液を除去後、T

DMEMで細胞表面を洗浄し、10%ウシ胎児血清含有DMEMを1穴あたり2ml加え、5%CO₂存在下、37°Cにて2日間培養した。

【0033】

2%ウシフィブリノーゲン（マイルス社）、0.5%アガロース、1mM塩化カルシウムを含む50mMリン酸緩衝液（pH 7.4）10mlに10単位のヒトトロンビン（持田製薬）を加え、直径9cmのプレート中で固化させ、フィブリンプレートを調製した。トランスフェクションしたCOS7細胞の培養上清10μlをフィブリンプレートに載せ、37°C 15時間インキュベートした。得られた溶解円の直径をウロキナーゼ活性の指標とした。表2に、各クローンからcDNA断片を切り出すのに用いた制限酵素部位、pSSD3を切斷するのに用いた制限酵素部位、溶解円の有無を示した。対照としてpSSD3を用いたもの以外は、いずれのサンプルも溶解円を形成し、ウロキナーゼが培地中に分泌されたことが確認された。すなわち、いずれのcDNA断片も分泌シグナル配列として機能するアミノ酸配列をコードしていることが示された。

【0034】

【表2】

表2

HP番号	制限酵素部位		溶解円
	cDNA*	ベクター	
HP00658	Hind III (K)	Sma I	+
HP00714	Pvu II	PmaC I	+
HP00876	Nco I (K)	PmaC I	+
HP01134	PmaC I	PmaC I	+
HP01269	Pvu II	PmaC I	+
HP10298	Hind III (K)	PmaC I	+

H P 1 0 3 6 8 E c o R V	P m a C I	+
p K A 1 - U P A		+
p S S D 3		-

* (K) は、制限酵素切斷後、クレノウ処理をすることを意味する。

【0035】

(6) インビトロ翻訳による蛋白質合成

本発明のcDNAを有するプラスミドベクターを用いて、T_NTウサギ網状赤血球溶解物キット（プロメガ社製）によるインビトロ転写／翻訳を行なった。この際 [³⁵S] メチオニンを添加し、発現産物をラジオアイソトープでラベルした。いずれの反応もキットに付属のプロトコールに従って行なった。プラスミド2 μ gを、T_NTウサギ網状赤血球溶解物12.5 μ l、緩衝液（キットに付属）0.5 μ l、アミノ酸混合液（Metを含まない）2 μ l、 [³⁵S] メチオニン（アマーシャム社）2 μ l（0.37MBq/ μ l）、T7 RNAポリメラーゼ0.5 μ l、RNasin 20Uを含む総量25 μ lの反応液中で30°Cで90分間反応させた。また、膜系存在下の実験は、この反応系に、イヌ臍臓ミクロソーム画分（プロメガ）2.5 μ lを添加して行った。反応液3 μ lにSDSサンプリングバッファー（125mMトリス塩酸緩衝液、pH6.8、120mM2-メルカプトエタノール、2% SDS溶液、0.025%プロモフェノールブルー、20%グリセロール）2 μ lを加え、95°C3分間加熱処理した後、SDS-ポリアクリルアミドゲル電気泳動にかけた。オートラジオグラフィーを行ない、翻訳産物の分子量を求めた。表3に、各クローンから得られた膜ミクロソーム系の存在下／非存在下におけるインビトロ翻訳産物の分子量の大きさを、cDNAのORFがコードしている蛋白質の分子量の計算値と共に示した。

【0036】

【表3】

表3

配列	H P 番号	計算値 (Da)	インビトロ翻訳産物 (kDa)
----	--------	----------	-----------------

番号		膜系無添加	膜系添加*
1	HP00658	17, 037	18
2	HP00714	37, 106	47
3	HP00876	18, 230	18
4	HP01134	42, 947	42
5	HP10029	18, 894	21
6	HP10189	9, 113	12
7	HP10269	129, 572	130
8	HP10298	13, 161	16
9	HP10368	19, 979	19

* - は未検討。

【0037】

(7) クローン例

<HP00658> (配列番号1、10、19)

ヒト繊維肉腫細胞株HT-1080 cDNAライブラリーから得られたクローンHP00658のcDNAインサートの全塩基配列を決定したところ、55 bpの5' 非翻訳領域、465 bpのORF、776 bpの3' 非翻訳領域からなる構造を有していた。ORFは154アミノ酸残基からなる蛋白質をコードしており、N末端に分泌シグナル配列と思われる疎水性領域が存在した。図3にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロファイルを示す。ORFがコードしているアミノ酸配列を用いてプロテインデータベースを検索したところ、N末端63アミノ酸残基はRANTES蛋白質(EMBLアクセスション番号M21121)と7番目の一アミノ酸残基(RANTESではマリゼー、大蛋白質ではマーラー)以外は完全に一致したが、64番目以降け

質は、それより長い154個アミノ酸残基からなっていた。インビトロ翻訳の結果、ORFから予想される分子量17, 037とほぼ一致する18 kDaの翻訳産

物が生成した。この際、ミクロソームを添加すると、分泌シグナル配列部分が切断除去されたと考えられる 16 kDa の産物が生成した。この結果は、pSSD 3 の結果と合わせて、本蛋白質が分泌シグナル配列を有することを確証する。分泌シグナル配列切断部位予測法である (-3, -1) 規則 [von Heijne, G., Nucl. Acid Res. 14: 4683-4690 (1986)] を適用すると、成熟蛋白質は 24 番目のセリンから始まると予想される。

【0038】

両者の塩基配列を比較すると、RANTES cDNA では、本 cDNA の 242 番目から 325 番目までの塩基配列が欠失していることがわかった。その結果、フレームシフトが起こり、サイズの異なる ORF が生成したと考えられる。他の領域でもいくつか変異が認められ、241 番目までは 97.7%、325 番目以降は 98.0% の相同性であった。RANTES は T 細胞特異的な蛋白質として得られたものであるが [Schall, T. J. et al., J. Immunol. 141: 1018-1025 (1988)]、本 cDNA は 繊維肉腫細胞から得られた。したがって、本蛋白質は RANTES とは異なる機能を有していると考えられる。

【0039】

また、本 cDNA の塩基配列を用いて GenBank を検索したが、EST の中に、90% 以上の相同性を有するものは見いだせなかった。

【0040】

<HP00714> (配列番号 2, 11, 20)

ヒト類表皮癌細胞株 KB cDNA ライブラリーから得られたクローン HP00714 の cDNA インサートの全塩基配列を決定したところ、56 bp の 5' 非翻訳領域、948 bp の ORF、2310 bp の 3' 非翻訳領域からなる構造を有していた。ORF は 315 アミノ酸残基からなる蛋白質をコードしており、N 末端に分泌シグナル配列と思われる疎水性領域が存在した。図 4 に Kyte-Doolittle の方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インビトロ翻訳の結果、ORF から予想される分子量 37,106 より大きい 47 kDa の翻訳産物が生成した。本蛋白質と類似性を有するヒトレティキュロ

カルビンの場合にも、SDS-PAGE上の翻訳産物のバンドが予想される分子量よりも10kDa程度大きめにでるので [Ozawa, M.、J. Bioc hem. 117: 1113-1119 (1995)]、この差異はこの蛋白質の物理化学性質に起因するものと考えられる。分泌シグナル配列切断部位予測法である(-3、-1)規則を適用すると、成熟蛋白質は20番目のリジンから始まる予想される。本蛋白質は、C末端に小胞体局在化シグナルモチーフ配列であるKDELに類似した配列HDEFを有しているので、小胞体に存在している可能性がある。

【0041】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ヒトレティキュロカルビン (GenBank アクセション番号D42073) と類似性を有していた。表4に、本発明のヒト蛋白質 (HP) とヒトレティキュロカルビン (RC) のアミノ酸配列の比較を示す。-はギャップを、*は本発明の蛋白質と同一アミノ酸残基を、.は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。両者は、60.5%の相同意を有していた。

【0042】

【表4】

表4

HP MDLRQFLMCLSLCTAFLSKPTEKKDR-VHHEPQLSDKVHNDAQSFYDH

.*. * * . . *** . *.* * . . . * . . . * . . ***.***

RC MARGGRGRRGLALGLLLALVLAPRVLRAKPTVRKERVVRPDSELGERPPEDNQSFQYDH

HP DAFLGAEAEAKTFDQLTPEESKERLGKIVSKIDGDKDGFVTVDDELKDWIKFFAQKRWIYEDV

.**** * ..*****.*****.**.* ****. .***.*** .***.* .*

RC EAFLGKEDSKTFDQLTPDESKERLGKIVDRIDNDGDFVTTEELKTWIKRVQKRYIFDNV

HP FRRQKQCHDIAXEDCI VCKEEVVVATVQVVI DDD DDDDDCEVVKOMMVRDERRRKKMADK

RC AKVWKDYDRDKDDKISWEFYKQATYGYYLGNPAEFHDSSDHHTFKMLPRDERKFKAADL

HP DGDLIATKEEFTAFLHPEEYDYMKDIVVQETMEDIDKNADGFIDLEEYIGDMYSHDGNTD

[0 0 4 3]

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同意し、かつ開始コドンから含んでいるもの（例えば、アクセション番号F3872）が存在したが、この配列から本蛋白質を予想することはできない。

【0 0 4 4】

レティキュロカルビンは、小胞体の膜表面に局在している蛋白質であり、小胞体内に分泌されてきた蛋白質のフォールディングなどに関与していると考えられている。したがって、本発明の蛋白質も、組換え蛋白質のフォールディングプロセスへの応用が考えられる。

[0 0 4 5]

<HP00876> (配列番号3、12、21)

ヒト胃癌cDNAライブラリーから得られたクローンHP00876のcDNAインサートの全塩基配列を決定したところ、146bpの5'非翻訳領域、477bpのORF、529bpの3'非翻訳領域からなる構造を有していた。ORFは158アミノ酸残基からなる蛋白質をコードしており、N末端に分泌シグナル配列と思われる疎水性領域が存在した。図5にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インビトロ翻訳の結果、ORFから予想される分子量18,230にほぼ一致する18kDaの翻訳産物が生成した。この際、ミクロソームを添加すると、分泌シグナル配列

部分が切断除去されたと考えられる 16 kDa の産物が生成した。この結果は、PSSD3 の結果と合わせて、本蛋白質が分泌シグナルを有することを確証する。分泌シグナル配列切断部位予測法である (-3, -1) 規則を適用すると、成熟蛋白質は 18 番目のグリシンか 23 番目のアスパラギン酸から始まると予想される。

【0046】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、いくつかの C 型レクチンと類似性を有していた。その一例として、表 5 に、本発明のヒト蛋白質 (HP) とガラガラヘビ C 型レクチン (CL) (Swiss-PROT アクセション番号 P21963) のアミノ酸配列の比較を示す。- はギャップを、* は本発明の蛋白質と同一アミノ酸残基を、. は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。両者は、35.3% の相同性を有していた。

【0047】

【表 5】

表 5

HP	MASRSMRLLLLSCLAKTGVLDIIMRPSCAPGWFYHKSNCYGYFRKLRNWSDAELECQS
	. * . * . .. * * * .. * . * . * . * . .
CL	NNCPLDWLPNGLCYKIFNQLKTWEDAEMFCRK
HP	YNGAHLASILSLKEASTIAEYISGYQRSQ-PIWIGLHDPQKRQQWQWIDGAMYLYRSWS
	* * * * . . * . * . * . * . * . * . * . * .
CL	YKPGCHLASFHRYGESLEIAEYISDYHKGQENVWIGLRDKKDFSWEWTDRCSTDYLTD
HP	GKSMGG--NKH-CAEMSSNNNFLTWSNECNKRQHFLCKYRP
	. . * * . * . . * . . * . . * . .
CL	KNQPDHYQNKEFCVELVSLTGYRLWNDQVCESKDAFLCQCKF

また、本 C 型レクチンの塩基配列を用いて GenBank を検索したが、BLAST の中に、90% 以上の相同性を有するものは見いだせなかった。

【0049】

プラスミドpHP00876を1 μ gを20単位のPvuIIで消化した後、1%アガロースゲル電気泳動にかけ、約700bpのDNA断片をゲルから切り出した。次いでpET-21a (Novagen社) 1 μ gを、20単位のNheIで消化した後クレノウ処理を行い、1%アガロースゲル電気泳動にかけ、5.4kbpのDNA断片をゲルから切り出した。ベクター断片とcDNA断片をライゲーションキットにより連結後、大腸菌BL21 (DE3) (Novagen社) を形質転換した。形質転換体からプラスミドpET876を調製し、制限酵素切断地図により目的とする組換え体を確認した。本発現ベクターは、クローンHP00876がコードする蛋白質の第29番目のセリンから始まる蛋白質の前にメチオニン-アラニンが付加した蛋白質を発現する。

【0050】

pET876/BL21 (DE3) を100 μ g/mlアンピシリン含有LB培地5mlに懸濁し、37℃で振とう培養し、 A_{600} が約0.5になったときにイソプロピルチオガラクトシドを1mMになるように添加した。さらに37℃で6時間培養後、遠心によって集菌し、菌体をアミロースカラム用カラム緩衝液 (10mMトリス塩酸、pH7.4、200mM NaCl、1mM EDTA) 25mlに懸濁した。この溶液を超音波処理後、不溶画分をSDS-ポリアクリルアミド電気泳動にかけたところ、約14kDaの位置に本ベクターの発現に由来するバンドが認められた。

【0051】

レクチンは、糖鎖を認識して結合するので、糖鎖検出試薬や糖蛋白質精製用アフィニティー担体としての用途がある。また、細胞外に分泌されたレクチンは細胞間情報伝達においても重要な役割を果たしているので、医薬としての用途もある。

【0052】

<HP01134> (配列番号4、13、22)

ヒト肝臓cDNAライブラリーから得られたクローンHP01134のcDNAインサートの全塩基配列を決定したところ、116bpの5'非翻訳領域、1

131 bp の ORF、502 bp の 3' 非翻訳領域からなる構造を有していた。ORF は 376 アミノ酸残基からなる蛋白質をコードしており、N 末端に分泌シグナル配列と思われる疎水性領域が存在した。図 6 に Kyte-Doolittle の方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インビトロ 翻訳の結果、ORF から予想される分子量 42,947 とほぼ一致する 42 kDa の翻訳産物が生成した。この際、ミクロソームを添加すると、分泌後 N-グリコシレーションによって糖鎖が付加されたと考えられる 49 kDa の産物が生成した。なお、この蛋白質のアミノ酸配列の中には、N-グリコシレーションが起こる可能性がある部位が 4箇所 (91 番目 Asn-Gly-Thr, 167 番目 Asn-Glu-Thr, 263 番目 Asn-Thr-Ser, 272 番目 Asn-Lys-Thr) 存在する。以上の結果は、PSSD3 の結果と合わせて、本蛋白質が分泌シグナルを有することを確認する。分泌シグナル配列切断部位予測法である (-3, -1) 規則を適用すると、成熟蛋白質は 17 番目のアラニンか 18 番目のバリンから始まると予想される。

【0053】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、いくつかのシステインプロテイナーゼと類似性を有していた。その一例として、表 6 に、本発明のヒト蛋白質 (HP) とミカンシステインプロテイナーゼ (CP) (GenBank アクセション番号 Z47793) のアミノ酸配列の比較を示す。- はギャップを、* は本発明の蛋白質と同一アミノ酸残基を、. は本発明の蛋白質と類似アミノ酸残基をそれぞれ表す。両者は、N 末端側 286 アミノ酸残基の領域で 49% の相同性を有していた。本蛋白質は、分泌された蛋白質のプロセシングに関与していると考えられる。

【0054】

【表 6】

表 6

HP

MIVKAVAVLSSVAVAGAVVADIPPELGKII

* * * .. * ..

CP MTRLASGVLITLLVALAGIADGSRDIAGDILKLPSEAYRFFHNGGGAKVNDDDSVGTR
HP WVVIVAGSNGWYNRHQADACHAYQI IHRNGIPDEQIVVMMYDDIAYSEDNPTPGIVINR

...*****..***** * * * .. * .. * * * .. * .. * .. *

CP WAVLLAGSNGFWNYRHQADICHAYQLLRKGGLKDENII VFMYDDIAFNEENPRPGVI INH
HP PNGTDVYQGVPKDYTGEDVTPQNFLAVLRGDAEAVKGIGSGKVLKSGPQDHVFIYFTDHG

..***.***** ..*.*. * ..*..* * * * ..*..* ..* ..* ..*

CP PHGDDVYKGVPKDYTGEDVTVEKFFAVVLGNKTALTG-GSGKVVDSGPNDHIFIFYSDHG
HP STGILVFPNED-LHVKDLNETIHYMYKHKMYRKMVFYIEACESGSMMN-HLPDNINVYAT

..*.** * ..*..*****... * ..*..*

CP GPGVLGMPTSRYIYADELIDVLKKKHASGNYKSLVFYLEACESGSIFEGLLLEGLNIYAT
HP TAANPRESSYACYY---DEKRSTY---LGDWYSVNWMEDSDVEDLTKEHLHQYHLVKS

.*. *.. *. ..*..* *** * ..*****...* ..*..* ..*

CP TASNAEESSWGTYCPGEIPGPPPEYSTCLGDLYSIAWMEDSDIHNLRTETLHQYELVKT
HP HT----NTSHVMQYGNKTISTMKVMQFQGMKRKASSPVPLPPVTHLDLTPSPDVPLTIM

.* ..*****. ..* .. . *

CP RTASYN SYGSHVMQYGDIGLSKNNLFTYLTNPANDNYTFVDENSLRPASKAVNQRDADL

【0055】

また、本cDNAの塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同意を有するもの（例えば、アクセション番号FO1300）が存在したが、いずれも本cDNAより短く、開始コドンから含んでいるものは見いだせなかった。

【0056】

細胞外に分泌されるプロテアーゼは、各種生理機能を有しているので、医薬としての用途を有している。また、限定分解による蛋白質の構造解析など、研究用試薬としても用いられている。

【0057】

<HP10029>（配列番号5、14、23）

ヒト類表皮癌細胞株KB cDNAライブラリーから得られたクローンHP10

029のcDNAインサートの全塩基配列を決定したところ、8bpの5'非翻訳領域、522bpのORF、458bpの3'非翻訳領域からなる構造を有していた。ORFは173アミノ酸残基からなる蛋白質をコードしており、N末端に分泌シグナル配列と思われる疎水性領域が存在した。図7にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インピトロ翻訳の結果、ORFから予想される分子量18,894とほぼ一致する21kDaの翻訳産物が生成した。この際、ミクロソームを添加すると、分泌シグナル配列部分が切断除去されたと考えられる18kDaの産物が生成した。この結果は、pSSD3の結果と合わせて、本蛋白質が分泌シグナル配列を有することを確認する。分泌シグナル配列切断部位予測法である(-3, -1)規則を適用すると、成熟蛋白質は32番目のバリンから始まると予想される。本蛋白質は、C末端に小胞体局在化シグナルモチーフ配列であるKDELに類似した配列RTELを有しているので、小胞体に存在している可能性がある。

【0058】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、相同性を有するものは見いだせなかった。ただし、塩基配列を用いてGenBankを検索したところ、ESTの中に、90%以上の相同性を有するもの（例えば、アクセション番号H87021）が存在したが、いずれも本cDNAより短く、開始コドンから含んでいるものは見いだせなかった。

【0059】

<HP10189>（配列番号6、15、24）

ヒト類表皮癌細胞株KB cDNAライブラリーから得られたクローンHP10189のcDNAインサートの全塩基配列を決定したところ、101bpの5'非翻訳領域、222bpのORF、67bpの3'非翻訳領域からなる構造を有していた。ORFは73アミノ酸残基からなる蛋白質をコードしており、N末端に分泌シグナル配列と思われる疎水性領域が存在した。図8にKyte-Doolittle

インピトロ翻訳の結果、ORFから予想される分子量9,115とはは一致する10kDaの翻訳産物が生成した。分泌シグナル配列切断部位予測法である(-3

、-1) 規則を適用すると、成熟蛋白質は27番目のアラニンから始まると予想される。

【0060】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、相同性を有するものは見いだせなかった。ただし、塩基配列を用いてGenBankを検索したところ、ESTの中に、本cDNAと90%以上の相同性を有し、開始コドンから含んでいるもの（例えば、アクセション番号N56270）が存在したが、フレームシフトが起こっており、本cDNAと同じORFは見いだせなかった。

【0061】

<HP10269>（配列番号7、16、25）

ヒトリンホーマ細胞株U937 cDNAライブラリーから得られたクローンHP10269のcDNAインサートの全塩基配列を決定したところ、753bpの5' 非翻訳領域、3519bpのORF、395bpの3' 非翻訳領域からなる構造を有していた。ORFは1172アミノ酸残基からなる蛋白質をコードしており、N末端に分泌シグナル配列と思われる疎水性領域が存在した。図9にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロファイルを示す。インビトロ翻訳の結果、ORFから予想される分子量129,571とほぼ一致する130kDaの翻訳産物が生成した。分泌シグナル配列切断部位予測法である（-3、-1）規則を適用すると、成熟蛋白質は18番目のグルタミンから始まると予想される。

【0062】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、ラミニンSのB3鎖と類似性を有していた。表7に、本発明のヒト蛋白質（HP）とヒトラミニンSのB3鎖（B3）（GenBankアクセション番号L25541）のアミノ酸配列の比較を示す。

【0063】

【表7】

表7

アミノ酸残基番号	H P	B 3
1 2 4	G l n	A r g
2 6 9	P r o	欠失
3 8 8	P r o	A l a
4 2 6	G l n	A r g
4 2 7	G l y	A r g
4 3 9	A r g	欠失
4 4 1	A s p	G l u
6 0 3	A r g	P r o
8 1 5	G l y	A l a

【0064】

本cDNAの塩基配列とデータベース記載の塩基配列を比較すると、本cDNAの方が5'末端が600bp以上長く、データベース記載の塩基配列の5'末端81bpは、本cDNAの塩基配列と全く一致しない。従って、両者は異なるmRNAに由来するものである。

【0065】

ラミニンは、細胞外マトリックスとして、細胞の増殖や分化に深く関わっている。したがって、細胞培養添加剤などとして用いられている。

【0066】

<HP10298> (配列番号8、17、26)

ヒト胃癌cDNAライブラリーから得られたクローンHP10298のcDNAインサートの全塩基配列を決定したところ、137bpの5'非翻訳領域、3'コドンの一つ目～二つ目の5'非翻訳領域からかく構造を有している。

ナル配列と思われる疎水性領域が存在した。図10にKyte-Doolittle法の方法で求めた本蛋白質の疎水性／親水性プロファイルを示す。インビトロ

翻訳の結果、ORFから予想される分子量13, 161とほぼ一致する16kDaの翻訳産物が生成した。分泌シグナル配列切断部位予測法である(-3, -1)規則を適用すると、成熟蛋白質は18番目のロイシンから始まると予想される。本蛋白質は、C末端約20アミノ酸残基が疎水性であることから、この部分を介して膜に結合している可能性もある。

【0067】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、相同性を有するものは見いだせなかった。ただし、塩基配列を用いてGenBankを検索したところ、ESTの中に、本cDNAと90%以上の相同性を有し、開始コドンから含んでいるもの（例えば、アクセション番号D78655）が存在したが、不明瞭な配列が多く、本cDNAと同じORFは見いだせなかった。

【0068】

<HP10368>（配列番号9、18、27）

ヒト胃癌cDNAライブラリーから得られたクローンHP10368のcDNAインサートの全塩基配列を決定したところ、72bpの5'非翻訳領域、528bpのORF、266bpの3'非翻訳領域からなる構造を有していた。ORFは175アミノ酸残基からなる蛋白質をコードしており、N末端に分泌シグナル配列と思われる疎水性領域が存在した。図11にKyte-Doolittleの方法で求めた本蛋白質の疎水性／親水性プロフィールを示す。インビトロ翻訳の結果、ORFから予想される分子量19, 979とほぼ一致する20kDaの翻訳産物が生成した。この際、ミクロソームを添加すると、分泌シグナル配列部分が切断除去されたと考えられる19kDaの産物が生成した。この結果は、pSSD3の結果と合わせて、本蛋白質が分泌シグナルを有することを確認する。分泌シグナル配列切断部位予測法である(-3, -1)規則を適用すると、成熟蛋白質は19番目のロイシンか21番目のアルギニンから始まると予想される。本蛋白質は、C末端に小胞体局在化シグナルモチーフ配列であるKDELに類似した配列KTELを有しているので、小胞体に存在している可能性がある。

【0069】

本蛋白質のアミノ酸配列を用いてプロテインデータベースを検索したところ、相同性を有するものは見いだせなかった。ただし、塩基配列を用いてGenBankを検索したところ、ESTの中に、本cDNAと90%以上の相同性を有し、開始コドンから含んでいるもの（例えば、アクセション番号T86663）が存在したが、不明瞭な配列が多く、本cDNAと同じORFは見いだせなかった。

【0070】

【発明の効果】

本発明は分泌シグナル配列を有するヒト蛋白質、およびそれをコードしているcDNAを提供する。本発明の蛋白質は、いずれも細胞外に分泌され、細胞外液中あるいは細胞膜表面に存在するので、細胞の増殖や分化を制御している蛋白質と考えられる。したがって、本発明の蛋白質は、医薬品として、あるいは該蛋白質に対する抗体を作製するための抗原として用いることができる。また、該DNAを用いることにより、該蛋白質を大量に発現することができる。

【0071】

【配列表】

配列番号：1

配列の長さ：154

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：纖維肉腫

セルライン：HT-1080

クローン名：HP00658

配列

Met	Lys	Val	Ser	Ala	Ala	Ala	Leu	Ala	Val	Ile	Leu	Ile	Ala	Thr	Ala
1										10					15
Leu	Cys	Ala	Pro	Ala	Ser	Ala	Ser	Pro	Tyr	Ser	Ser	Asp	Thr	Thr	Pro
										25					30
Cys	Cys	Phe	Ala	Tyr	Ile	Ala	Arg	Pro	Leu	Pro	Arg	Ala	His	Ile	Lys
										35					40
Glu	Tyr	Phe	Tyr	Thr	Ser	Gly	Lys	Cys	Ser	Asn	Pro	Ala	Val	Val	His
										50					55
Arg	Ser	Arg	Met	Pro	Lys	Arg	Glu	Gly	Gln	Gln	Val	Trp	Gln	Asp	Phe
										65					70
Leu	Tyr	Asp	Ser	Arg	Leu	Asn	Lys	Gly	Lys	Leu	Cys	His	Pro	Lys	Glu
										85					90
Pro	Pro	Ser	Val	Cys	Gln	Pro	Arg	Glu	Glu	Met	Gly	Ser	Gly	Val	His
										100					105
Gln	Leu	Phe	Gly	Asp	Glu	Leu	Gly	Trp	Arg	Val	Leu	Glu	Pro	Glu	Leu
										115					120
															125

特平 8-243060

Thr Gln Ile Cys Leu Phe Leu Leu Ala Leu Val Leu Ala Trp Glu Ala

130 135 140

Ser Pro His Tyr Pro Thr Pro Pro Ala Pro

145 150

【0072】

特平 8-243060

配列番号：2

配列の長さ：315

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル: No

起源:

生物名: ホモ=サピエンス

細胞の種類: 類表皮癌

セルライン: KB

クローン名: HP00714

配列

Met	Asp	Leu	Arg	Gln	Phe	Leu	Met	Cys	Leu	Ser	Leu	Cys	Thr	Ala	Phe
1							10							15	
Ala	Leu	Ser	Lys	Pro	Thr	Glu	Lys	Lys	Asp	Arg	Val	His	His	Glu	Pro
							25							30	
Gln	Leu	Ser	Asp	Lys	Val	His	Asn	Asp	Ala	Gln	Ser	Phe	Asp	Tyr	Asp
							35							45	
His	Asp	Ala	Phe	Leu	Gly	Ala	Glu	Glu	Ala	Lys	Thr	Phe	Asp	Gln	Leu
							50							60	
Thr	Pro	Glu	Glu	Ser	Lys	Glu	Arg	Leu	Gly	Lys	Ile	Val	Ser	Lys	Ile
							65							80	
Asp	Gly	Asp	Lys	Asp	Gly	Phe	Val	Thr	Val	Asp	Glu	Leu	Lys	Asp	Trp
							85							95	
Ile	Lys	Phe	Ala	Gln	Lys	Arg	Trp	Ile	Tyr	Glu	Asp	Val	Glu	Arg	Gln
							100							110	
Trp	Lys	Gly	His	Asp	Leu	Asn	Glu	Asp	Gly	Leu	Val	Ser	Trp	Glu	Glu
							115							125	
Tyr	Lys	Asn	Ala	Thr	Tyr	Gly	Tyr	Val	Leu	Asp	Asp	Pro	Asp	Pro	Asp
							130							140	
Ile	Cys	Phe	Asp	Tyr	Lys	Cys	Met	Met	Val	Arg	Asp	Cys	Arg	Arg	Phe
Lys	Met	Ala	Asp	Lys	Asp	Gly	Asp	Leu	Ile	Ala	Thr	Lys	Glu	Glu	Phe
							165							175	

Thr Ala Phe Leu His Pro Glu Glu Tyr Asp Tyr Met Lys Asp Ile Val
 180 185 190
 Val Gln Glu Thr Met Glu Asp Ile Asp Lys Asn Ala Asp Gly Phe Ile
 195 200 205
 Asp Leu Glu Glu Tyr Ile Gly Asp Met Tyr Ser His Asp Gly Asn Thr
 210 215 220
 Asp Glu Pro Glu Trp Val Lys Thr Glu Arg Glu Gln Phe Val Glu Phe
 225 230 235 240
 Arg Asp Lys Asn Arg Asp Gly Lys Met Asp Lys Glu Glu Thr Lys Asp
 245 250 255
 Trp Ile Leu Pro Ser Asp Tyr Asp His Ala Glu Ala Glu Ala Arg His
 260 265 270
 Leu Val Tyr Glu Ser Asp Gln Asn Lys Asp Gly Lys Leu Thr Lys Glu
 275 280 285
 Glu Ile Val Asp Lys Tyr Asp Leu Phe Val Gly Ser Gln Ala Thr Asp
 290 295 300
 Phe Gly Glu Ala Leu Val Arg His Asp Glu Phe
 305 310 315

【0073】

特平 8-243060

Glu Cys Asn Lys Arg Gln His Phe Leu Cys Lys Tyr Arg Pro

145 150

155

【0074】

配列番号：4

配列の長さ：376

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：肝臓

クローン名：HP01134

配列

Met Val Trp Lys Val Ala Val Phe Leu Ser Val Ala Leu Gly Ile Gly

1 5 10 15

Ala Val Pro Ile Asp Asp Pro Glu Asp Gly Gly Lys His Trp Val Val

20 25 30

Ile Val Ala Gly Ser Asn Gly Trp Tyr Asn Tyr Arg His Gln Ala Asp

35 40 45

Ala Cys His Ala Tyr Gln Ile Ile His Arg Asn Gly Ile Pro Asp Glu

50 55 60

Gln Ile Val Val Met Met Tyr Asp Asp Ile Ala Tyr Ser Glu Asp Asn

65 70 75 80

Pro Thr Pro Gly Ile Val Ile Asn Arg Pro Asn Gly Thr Asp Val Tyr

85 90 95

Gln Gly Val Pro Lys Asp Tyr Thr Gly Glu Asp Val Thr Pro Gln Asn

100 105 110

Phe Leu Ala Val Leu Arg C1v Asp Ala C1u Ala Val Lys C1v Ile C1v

Ser Gly Lys Val Leu Lys Ser Gly Pro Gln Asp His Val Phe Ile Tyr

130 135 140

Phe Thr Asp His Gly Ser Thr Gly Ile Leu Val Phe Pro Asn Glu Asp
 145 150 155 160
 Leu His Val Lys Asp Leu Asn Glu Thr Ile His Tyr Met Tyr Lys His
 165 170 175
 Lys Met Tyr Arg Lys Met Val Phe Tyr Ile Glu Ala Cys Glu Ser Gly
 180 185 190
 Ser Met Met Asn His Leu Pro Asp Asn Ile Asn Val Tyr Ala Thr Thr
 195 200 205
 Ala Ala Asn Pro Arg Glu Ser Ser Tyr Ala Cys Tyr Tyr Asp Glu Lys
 210 215 220
 Arg Ser Thr Tyr Leu Gly Asp Trp Tyr Ser Val Asn Trp Met Glu Asp
 225 230 235 240
 Ser Asp Val Glu Asp Leu Thr Lys Glu Thr Leu His Lys Gln Tyr His
 245 250 255
 Leu Val Lys Ser His Thr Asn Thr Ser His Val Met Gln Tyr Gly Asn
 260 265 270
 Lys Thr Ile Ser Thr Met Lys Val Met Gln Phe Gln Gly Met Lys Arg
 275 280 285
 Lys Ala Ser Ser Pro Val Pro Leu Pro Pro Val Thr His Leu Asp Leu
 290 295 300
 Thr Pro Ser Pro Asp Val Pro Leu Thr Ile Met Lys Arg Lys Leu Met
 305 310 315 320
 Asn Thr Asn Asp Leu Glu Glu Ser Arg Gln Leu Thr Glu Glu Ile Gln
 325 330 335
 Arg His Leu Asp Tyr Glu Tyr Ala Leu Arg His Leu Tyr Val Leu Val
 340 345 350
 Asn Leu Cys Glu Lys Pro Tyr Pro Leu His Arg Ile Lys Leu Ser Met
 355 360 365
 Asp His Val Cys Leu Gly His Tyr

特平 8-243060

370

375

【0075】

配列番号：5

配列の長さ：173

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：類表皮癌

セルライン：KB

クローン名：HP10029

配列

Met	Ala	Ala	Pro	Ser	Gly	Gly	Trp	Asn	Gly	Val	Arg	Ala	Ser	Leu	Trp
1															
Ala	Ala	Leu	Leu	Leu	Gly	Ala	Val	Ala	Leu	Arg	Pro	Ala	Glu	Ala	Val
		20					25						30		
Ser	Glu	Pro	Thr	Thr	Val	Ala	Phe	Asp	Val	Arg	Pro	Gly	Gly	Val	Val
35						40						45			
His	Ser	Phe	Ser	His	Asn	Val	Gly	Pro	Gly	Asp	Lys	Tyr	Thr	Cys	Met
50						55					60				
Phe	Thr	Tyr	Ala	Ser	Gln	Gly	Gly	Thr	Asn	Glu	Gln	Trp	Gln	Met	Ser
65						70				75			80		
Leu	Gly	Thr	Ser	Glu	Asp	His	Gln	His	Phe	Thr	Cys	Thr	Ile	Trp	Arg
85							90						95		
Pro	Gln	Gly	Lys	Ser	Tyr	Leu	Tyr	Phe	Thr	Gln	Phe	Lys	Ala	Glu	Val
100							105					110			
Arg	Gly	Ala	Glu	Ile	Glu	Tyr	Ala	Met	Ala	Tyr	Ser	Lys	Ala	Ala	Phe
115							120					125			
Glu	Arg	Glu	Ser	Asp	Val	Pro	Leu	Lys	Thr	Glu	Glu	Phe	Glu	Val	Thr

特平 8-243060

130

135

140

Lys Thr Ala Val Ala His Arg Pro Gly Ala Phe Lys Ala Glu Leu Ser

145

150

155

160

Lys Leu Val Ile Val Ala Lys Ala Ser Arg Thr Glu Leu

165

170

【0076】

配列番号：6

配列の長さ：73

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：類表皮癌

セルライン：KB

クローン名：HP10189

配列

Met Gly Val Lys Leu Glu Ile Phe Arg Met Ile Ile Tyr Leu Thr Phe

1 5 10 15

Pro Val Ala Met Phe Trp Val Ser Asn Gln Ala Glu Trp Phe Glu Asp

20 25 30

Asp Val Ile Gln Arg Lys Arg Glu Leu Trp Pro Pro Glu Lys Leu Gln

35 40 45

Glu Ile Glu Glu Phe Lys Glu Arg Leu Arg Lys Arg Arg Glu Glu Lys

50 55 60

Leu Leu Arg Asp Ala Gln Gln Asn Ser

65 70

【0077】

配列番号：7

配列の長さ：1172

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：組織球リンホーマ

セルライン：U937

クローン名：HP10269

配列

Met Arg Pro Phe Phe Leu Leu Cys Phe Ala Leu Pro Gly Leu Leu His

1

5

10

15

Ala Gln Gln Ala Cys Ser Arg Gly Ala Cys Tyr Pro Pro Val Gly Asp

20

25

30

Leu Leu Val Gly Arg Thr Arg Phe Leu Arg Ala Ser Ser Thr Cys Gly

35

40

45

Leu Thr Lys Pro Glu Thr Tyr Cys Thr Gln Tyr Gly Glu Trp Gln Met

50

55

60

Lys Cys Cys Lys Cys Asp Ser Arg Gln Pro His Asn Tyr Tyr Ser His

65

70

75

80

Arg Val Glu Asn Val Ala Ser Ser Gly Pro Met Arg Trp Trp Gln

85

90

95

Ser Gln Asn Asp Val Asn Pro Val Ser Leu Gln Leu Asp Leu Asp Arg

100

105

110

Ala Gly Met Leu Ile Glu Arg Ser Ser Asp Phe Gly Lys Thr Trp Arg

130	135	140
Val Tyr Gln Tyr Leu Ala Ala Asp Cys Thr Ser Thr Phe Pro Arg Val		
145	150	155
Arg Gln Gly Arg Pro Gln Ser Trp Gln Asp Val Arg Cys Gln Ser Leu		
165	170	175
Pro Gln Arg Pro Asn Ala Arg Leu Asn Gly Gly Lys Val Gln Leu Asn		
180	185	190
Leu Met Asp Leu Val Ser Gly Ile Pro Ala Thr Gln Ser Gln Lys Ile		
195	200	205
Gln Glu Val Gly Glu Ile Thr Asn Leu Arg Val Asn Phe Thr Arg Leu		
210	215	220
Ala Pro Val Pro Gln Arg Gly Tyr His Pro Pro Ser Ala Tyr Tyr Ala		
225	230	235
Val Ser Gln Leu Arg Leu Gln Gly Ser Cys Phe Cys His Gly His Ala		
245	250	255
Asp Arg Cys Ala Pro Lys Pro Gly Ala Ser Ala Gly Pro Ser Thr Ala		
260	265	270
Val Gln Val His Asp Val Cys Val Cys Gln His Asn Thr Ala Gly Pro		
275	280	285
Asn Cys Glu Arg Cys Ala Pro Phe Tyr Asn Asn Arg Pro Trp Arg Pro		
290	295	300
Ala Glu Gly Gln Asp Ala His Glu Cys Gln Arg Cys Asp Cys Asn Gly		
305	310	315
His Ser Glu Thr Cys His Phe Asp Pro Ala Val Phe Ala Ala Ser Gln		
325	330	335
Gly Ala Tyr Gly Gly Val Cys Asp Asn Cys Arg Asp His Thr Glu Gly		
340	345	350
Lys Asn Cys Glu Arg Cys Gln Leu His Tyr Phe Arg Asn Arg Arg Pro		
355	360	365

Gly Ala Ser Ile Gln Glu Thr Cys Ile Ser Cys Glu Cys Asp Pro Asp
 370 375 380
 Gly Ala Val Pro Gly Ala Pro Cys Asp Pro Val Thr Gly Gln Cys Val
 385 390 395 400
 Cys Lys Glu His Val Gln Gly Glu Arg Cys Asp Leu Cys Lys Pro Gly
 405 410 415
 Phe Thr Gly Leu Thr Tyr Ala Asn Pro Gln Gly Cys His Arg Cys Asp
 420 425 430
 Cys Asn Ile Leu Gly Ser Arg Arg Asp Met Pro Cys Asp Glu Glu Ser
 435 440 445
 Gly Arg Cys Leu Cys Leu Pro Asn Val Val Gly Pro Lys Cys Asp Gln
 450 455 460
 Cys Ala Pro Tyr His Trp Lys Leu Ala Ser Gly Gln Gly Cys Glu Pro
 465 470 475 480
 Cys Ala Cys Asp Pro His Asn Ser Leu Ser Pro Gln Cys Asn Gln Phe
 485 490 495
 Thr Gly Gln Cys Pro Cys Arg Glu Gly Phe Gly Gly Leu Met Cys Ser
 500 505 510
 Ala Ala Ala Ile Arg Gln Cys Pro Asp Arg Thr Tyr Gly Asp Val Ala
 515 520 525
 Thr Gly Cys Arg Ala Cys Asp Cys Asp Phe Arg Gly Thr Glu Gly Pro
 530 535 540
 Gly Cys Asp Lys Ala Ser Gly Arg Cys Leu Cys Arg Pro Gly Leu Thr
 545 550 555 560
 Gly Pro Arg Cys Asp Gln Cys Gln Arg Gly Tyr Cys Asn Arg Tyr Pro
 565 570 575
 580 585 590
 Arg Glu Gln Ala Leu Arg Phe Gly Arg Leu Arg Asn Ala Thr Ala Ser

595	600	605
Leu Trp Ser Gly Pro Gly Leu Glu Asp Arg Gly Leu Ala Ser Arg Ile		
610	615	620
Leu Asp Ala Lys Ser Lys Ile Glu Gln Ile Arg Ala Val Leu Ser Ser		
625	630	635
Pro Ala Val Thr Glu Gln Glu Val Ala Gln Val Ala Ser Ala Ile Leu		
645	650	655
Ser Leu Arg Arg Thr Leu Gln Gly Leu Gln Leu Asp Leu Pro Leu Glu		
660	665	670
Glu Glu Thr Leu Ser Leu Pro Arg Asp Leu Glu Ser Leu Asp Arg Ser		
675	680	685
Phe Asn Gly Leu Leu Thr Met Tyr Gln Arg Lys Arg Glu Gln Phe Glu		
690	695	700
Lys Ile Ser Ser Ala Asp Pro Ser Gly Ala Phe Arg Met Leu Ser Thr		
705	710	715
Ala Tyr Glu Gln Ser Ala Gln Ala Gln Gln Val Ser Asp Ser Ser		
725	730	735
Arg Leu Leu Asp Gln Leu Arg Asp Ser Arg Arg Glu Ala Glu Arg Leu		
740	745	750
Val Arg Gln Ala Gly Gly Gly Gly Thr Gly Ser Pro Lys Leu Val		
755	760	765
Ala Leu Arg Leu Glu Met Ser Ser Leu Pro Asp Leu Thr Pro Thr Phe		
770	775	780
Asn Lys Leu Cys Gly Asn Ser Arg Gln Met Ala Cys Thr Pro Ile Ser		
785	790	795
Cys Pro Gly Glu Leu Cys Pro Gln Asp Asn Gly Thr Ala Cys Gly Ser		
805	810	815
Arg Cys Arg Gly Val Leu Pro Arg Ala Gly Gly Ala Phe Leu Met Ala		
820	825	830

Gly Gln Val Ala Glu Gln Leu Arg Gly Phe Asn Ala Gln Leu Gln Arg
 835 840 845
 Thr Arg Gln Met Ile Arg Ala Ala Glu Glu Ser Ala Ser Gln Ile Gln
 850 855 860
 Ser Ser Ala Gln Arg Leu Glu Thr Gln Val Ser Ala Ser Arg Ser Gln
 865 870 875 880
 Met Glu Glu Asp Val Arg Arg Thr Arg Leu Leu Ile Gln Gln Val Arg
 885 890 895
 Asp Phe Leu Thr Asp Pro Asp Thr Asp Ala Ala Thr Ile Gln Glu Val
 900 905 910
 Ser Glu Ala Val Leu Ala Leu Trp Leu Pro Thr Asp Ser Ala Thr Val
 915 920 925
 Leu Gln Lys Met Asn Glu Ile Gln Ala Ile Ala Ala Arg Leu Pro Asn
 930 935 940
 Val Asp Leu Val Leu Ser Gln Thr Lys Gln Asp Ile Ala Arg Ala Arg
 945 950 955 960
 Arg Leu Gln Ala Glu Ala Glu Ala Arg Ser Arg Ala His Ala Val
 965 970 975
 Glu Gly Gln Val Glu Asp Val Val Gly Asn Leu Arg Gln Gly Thr Val
 980 985 990
 Ala Leu Gln Glu Ala Gln Asp Thr Met Gln Gly Thr Ser Arg Ser Leu
 995 1000 1005
 Arg Leu Ile Gln Asp Arg Val Ala Glu Val Gln Gln Val Leu Arg Pro
 1010 1015 1020
 Ala Glu Lys Leu Val Thr Ser Met Thr Lys Gln Leu Gly Asp Phe Trp
 1025 1030 1035 1040
 Ala Val Gln Ala Gln Gln Leu Ala Glu Gly Ala Ser Glu Gln Ala Leu

1060 1065 1070
Ser Ala Gln Glu Gly Phe Glu Arg Ile Lys Gln Lys Tyr Ala Glu Leu
1075 1080 1085
Lys Asp Arg Leu Gly Gln Ser Ser Met Leu Gly Glu Gln Gly Ala Arg
1090 1095 1100
Ile Gln Ser Val Lys Thr Glu Ala Glu Glu Leu Phe Gly Glu Thr Met
1105 1110 1115 1120
Glu Met Met Asp Arg Met Lys Asp Met Glu Leu Glu Leu Leu Arg Gly
1125 1130 1135
Ser Gln Ala Ile Met Leu Arg Ser Ala Asp Leu Thr Gly Leu Glu Lys
1140 1145 1150
Arg Val Glu Gln Ile Arg Asp His Ile Asn Gly Arg Val Leu Tyr Tyr
1155 1160 1165
Ala Thr Cys Lys
1170
【0078】

配列番号：8

配列の長さ：122

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：HP10298

配列

Met	Gly	Leu	Leu	Leu	Leu	Val	Pro	Leu	Leu	Leu	Leu	Pro	Gly	Ser	Tyr
1		5				10						15			
Gly	Leu	Pro	Phe	Tyr	Asn	Gly	Phe	Tyr	Tyr	Ser	Asn	Ser	Ala	Asn	Asp
					20			25				30			
Gln	Asn	Leu	Gly	Asn	Gly	His	Gly	Lys	Asp	Leu	Leu	Asn	Gly	Val	Lys
						35		40				45			
Leu	Val	Val	Glu	Thr	Pro	Glu	Glu	Thr	Leu	Phe	Thr	Arg	Ile	Leu	Thr
					50		55				60				
Val	Gly	Pro	Gln	Ser	Leu	Gly	Ser	Glu	Ala	Leu	Ala	Ser	Pro	Thr	Arg
					65		70			75			80		
Arg	Ala	Ala	Cys	Thr	Val	Phe	Thr	Ala	Thr	Ala	Ser	Thr	Arg	Thr	Trp
					85			90				95			
Gly	Pro	Pro	Leu	Pro	His	Ser	Leu	Thr	Gly	Cys	Val	Phe	Ile	Glu	Trp
					100			105			110				
Phe	Val	Phe	Pro	Cys	Cys	Ile	Cys	Ile	Cys	Ile	Cys	Pro	Phe		

[333]

配列番号：9

配列の長さ：175

配列の型：アミノ酸

トポロジー：直鎖状

配列の種類：蛋白質

ハイポセティカル：No

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：HP10368

配列

Met	Glu	Lys	Ile	Pro	Val	Ser	Ala	Phe	Leu	Leu	Leu	Val	Ala	Leu	Ser
1															
Tyr	Thr	Leu	Ala	Arg	Asp	Thr	Thr	Val	Lys	Pro	Gly	Ala	Lys	Lys	Asp
20															
Thr	Lys	Asp	Ser	Arg	Pro	Lys	Leu	Pro	Gln	Thr	Leu	Ser	Arg	Gly	Trp
35															
Gly	Asp	Gln	Leu	Ile	Trp	Thr	Gln	Thr	Tyr	Glu	Glu	Ala	Leu	Tyr	Lys
50															
Ser	Lys	Thr	Ser	Asn	Lys	Pro	Leu	Met	Ile	Ile	His	His	Leu	Asp	Glu
65															
Cys	Pro	His	Ser	Gln	Ala	Leu	Lys	Lys	Val	Phe	Ala	Glu	Asn	Lys	Glu
85															
Ile	Gln	Lys	Leu	Ala	Glu	Gln	Phe	Val	Leu	Leu	Asn	Leu	Val	Tyr	Glu
100															
Thr	Thr	Asp	Lys	His	Leu	Ser	Pro	Asp	Gly	Gln	Tyr	Val	Pro	Arg	Ile
115															
Met	Phe	Val	Asp	Pro	Ser	Leu	Thr	Val	Arg	Ala	Asp	Ile	Thr	Gly	Arg
130															
135															
140															

特平 8-243060

Tyr Ser Asn Arg Leu Tyr Ala Tyr Glu Pro Ala Asp Thr Ala Leu Leu

145 150 155 160

Leu Asp Asn Met Lys Lys Ala Leu Lys Leu Leu Lys Thr Glu Leu

165 170 175

【0080】

配列番号：10

配列の長さ：462

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：纖維肉腫

セルライン：HT-1080

クローン名：HP00658

配列

ATGAAGGTCT CCGCGGCAGC CCTCGCTGTC ATCCTCATTG CTACTGCCCT CTGCGCTCCT	60
GCATCTGCCT CCCCATATTC CTCGGACACC ACACCCCTGCT GCTTGCCTA CATTGCCCGC	120
CCACTGCCCC GTGCCCACAT CAAGGAGTAT TTCTACACCA GTGGCAAGTG CTCCAACCCA	180
GCAGTCGTCC ACAGGTCAAG GATGCCAAAG AGAGAGGGAC AGCAAGTCTG GCAGGATTTC	240
CTGTATGACT CCCGGCTGAA CAAGGGCAAG CTTTGTCAAC CGAAAGAACCGCCAAAGTG	300
TGCCAACCCA GAGAAGAAAT GGGTTCGGGA GTACATCAAC TCTTGAGA TGAGCTAGGA	360
TGGAGAGTCC TTGAACCTGA ACTTACACAA ATTTGCCTGT TTCTGCTTGC TCTTGTCTA	420
GCTTGGGAGG CTTCCCTCA CTATCCTACC CCACCCGCTC CT	462

【0081】

配列番号：11

配列の長さ：945

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：類表皮癌

セルライン：KB

クローン名：HP00714

配列

ATGGACCTGC GACAGTTCT TATGTGCCTG TCCCTGTGCA CAGCCTTGC CTTGAGCAAA	60
CCCACAGAAA AGAAGGACCG TGTACATCAT GAGCCTCAGC TCAGTGACAA GGTCACAAT	120
GATGCTCAGA GTTTGATTA TGACCATGAT GCCTTCTTGG GTGCTGAAGA AGCAAAGACC	180
TTTGATCAGC TGACACCAGA AGAGAGCAAG GAAAGGCTTG GAAAGATTGT AAGTAAAATA	240
GATGGCGACA AGGACGGGTT TGTCACTGTG GATGAGCTCA AAGACTGGAT TAAATTGCA	300
CAAAAGCGCT GGATTTACGA GGATGTAGAG CGACAGTGGAG AGGGGCATGA CCTCAATGAG	360
GACGGCCTCG TTTCTGGGA GGAGTATAAA AATGCCACCT ACGGCTACGT TTTAGATGAT	420
CCAGATCCTG ATGATGGATT TAACTATAAA CAGATGATGG TTAGAGATGA GCGGAGGTTT	480
AAAATGGCAG ACAAGGATGG AGACCTCATT GCCACCAAGG AGGAGTTCAC AGCTTCCCTG	540
CACCTGAGG AGTATGACTA CATGAAAGAT ATAGTAGTAC AGGAAACAAT GGAAGATATA	600
GATAAGAATG CTGATGGTTT CATTGATCTA GAAGAGTATA TTGGTGACAT GTACAGCCAT	660
GATGGGAATA CTGATGAGCC AGAATGGTA AAGACAGAGC GAGAGCAGTT TGTTGAGTTT	720
CGGGATAAGA ACCGTGATGG GAAGATGGAC AAGGAAGAGA CCAAAGACTG GATCCTTCCC	780
TGACACTATC ATCATGCCACA CCCACAAACCC ACCCACCTCC TCTATCAATC ACACCAAAAC	840

AGGGCAACAG AATTTGGGAA GGGCTTAAGTA CGGGAATGATG AGTTG	845
---	-----

【0082】

配列番号： 12

配列の長さ： 474

配列の型： 核酸

鎖の数： 二本鎖

トポロジー： 直鎖状

配列の種類： cDNA to mRNA

起源：

生物名： ホモ＝サピエンス

細胞の種類： 胃癌

クローン名： HP00876

配列

ATGGCTTCCA	GAAGCATGCG	GCTGCTCCTA	TTGCTGAGCT	GCCTGGCCAA	AACAGGAGTC	60
CTGGGTGATA	TCATCATGAG	ACCCAGCTGT	GCTCCTGGAT	GGTTTTACCA	CAAGTCCAAT	120
TGCTATGGTT	ACTTCAGGAA	GCTGAGGAAC	TGGTCTGATG	CCGAGCTCGA	GTGTCAGTCT	180
TACGGAAACG	GAGCCCACCT	GGCATCTATC	CTGAGTTAA	AGGAAGCCAG	CACCATAGCA	240
GAGTACATAA	GTGGCTATCA	GAGAAGCCAG	CCGATATGGA	TTGGCCTGCA	CGACCCACAG	300
AAGAGGCAGC	AGTGGCAGTG	GATTGATGGG	GCCATGTATC	TGTACAGATC	CTGGTCTGGC	360
AAGTCCATGG	GTGGGAACAA	GCACTGTGCT	GAGATGAGCT	CCAATAACAA	CTTTTTAACT	420
TGGAGCAGCA	ACGAATGCAA	CAAGCGCCAA	CACTTCCTGT	GCAAGTACCG	ACCA	474

【0083】

特平 8-243060

配列番号：13

配列の長さ：1128

配列の型：核酸

鎖の数： 二本鎖

トポロジー：直鎖状

配列の種類： cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：肝臓

クローン名：HP01134

配列

ATGGTTTGG	AAGTAGCTGT	ATTCCCTCAGT	GTGGCCCTGG	GCATTGGTGC	CGTCCCTATA	60
GATGATCCTG	AAGATGGAGG	CAAGCACTGG	GTGGTGATCG	TGGCAGGTT	AAATGGCTGG	120
TATAATTATA	GGCACCCAGGC	AGACCGCGTGC	CATGCCTACC	AGATCATTCA	CCGCAATGGG	180
ATTCCCTGACG	AACAGATCGT	TGTGATGATG	TACGATGACA	TTGCTTACTC	TGAAGACAAT	240
CCCACTCCAG	GAATTGTGAT	CAACAGGCC	AATGGCACAG	ATGTCTATCA	GGGAGTCCCG	300
AAGGACTACA	CTGGAGAGGA	TGTTACCCCA	CAAAATTCC	TTGCTGTGTT	GAGAGGCGAT	360
GCAGAAGCAG	TGAAGGGCAT	AGGATCCGGC	AAAGTCCTGA	AGAGTGGCCC	CCAGGATCAC	420
GTGTCATTT	ACTTCACTGA	CCATGGATCT	ACTGGAATAC	TGGTTTTCC	CAATGAAGAT	480
CTTCATGTAA	AGGACCTGAA	TGAGACCATC	CATTACATGT	ACAAACACAA	AATGTACCGA	540
AAGATGGTGT	TCTACATTGA	AGCCTGTGAG	TCTGGGTCCA	TGATGAACCA	CCTGCCGGAT	600
AACATCAATG	TTTATGCAAC	TACTGCTGCC	AACCCCAGAG	AGTCGTCTTA	CGCCTGTTAC	660
TATGATGAGA	AGAGGTCCAC	GTACCTGGGG	GACTGGTACA	GCGTCAACTG	GATGGAAGAC	720
TCGGACGTGG	AAGATCTGAC	TAAAGAGACC	CTGCACAAGC	AGTACCAACT	GGTAAAATCG	780
CACACCAACA	CCAGCCACGT	CATGCAGTAT	GGAAACAAAA	CAATCTCCAC	CATGAAAGTG	840
ATGCAGTTTC	AGGGTATGAA	ACGCAAAGCC	AGTTCTCCCG	TCCCCCTACC	TCCAGTCACA	900
CACCTTGACC	TCACCCCCAG	CCCTGATGTG	CCTCTCACCA	TCATGAAAAG	GAAACTGATG	960
AACACCAATG	ATCTGGAGGA	GTCCAGGCAG	CTCACGGAGG	AGATCCAGCG	GCATCTGGAT	1020
TACGAGTATG	CGTTGAGACA	TTTGTACGTG	CTGGTCAACC	TTTGTGAGAA	GCCGTATCCG	1080
CTTCACAGGA	TAAAATTGTC	CATGGACCAC	GTGTGCCCTTG	GTCACTAC		1128

【0084】

配列番号： 14

配列の長さ： 519

配列の型： 核酸

鎖の数： 二本鎖

トポロジー： 直鎖状

配列の種類： cDNA to mRNA

起源：

生物名： ホモ＝サピエンス

細胞の種類： 類表皮癌

セルライン： KB

クローン名： HP10029

配列

ATGGCGGCCG	CCAGCGGAGG	GTGGAACGGC	GTCCGCGCGA	GCTTGTGGGC	CGCGCTGCTC	60
CTAGGGGCCG	TGGCGCTGAG	GCCGGCGGAG	GCGGTGTCCG	AGCCCACGAC	CGTGGCGTT	120
GACGTGCGGC	CCGGCGGCCGT	CGTGCATTCC	TTCTCCCATA	ACGTGGGCC	GGGGGACAAA	180
TATACGTGTA	TGTTCACTTA	CGCCTCTCAA	GGAGGGACCA	ATGAGCAATG	GCAGATGAGT	240
CTGGGGACCA	GCGAAGACCA	CCAGCACTTC	ACCTGCACCA	TCTGGAGGCC	CCAGGGGAAG	300
TCCTATCTGT	ACTTCACACA	GTTCAAGGCA	GAGGTGCGGG	GCGCTGAGAT	TGAGTACGCC	360
ATGGCCTACT	CTAAAGCCGC	ATTGAAAGG	GAAAGTGATG	TCCCTCTGAA	AACTGAGGAA	420
TTTGAAGTGA	CCAAAACAGC	AGTGGCTCAC	AGGCCCGGGG	CATTCAAAGC	TGAGCTGTCC	480
AAGCTGGTGA	TTGTGGCCAA	GGCATCGCGC	ACTGAGCTG			519

【0085】

配列番号： 15

配列の長さ： 219

配列の型： 核酸

鎖の数： 二本鎖

トポロジー： 直鎖状

配列の種類： cDNA to mRNA

起源：

生物名： ホモ＝サピエンス

細胞の種類： 類表皮癌

セルライン： KB

クローン名： HP10189

配列

ATGGGGGTGA AGCTGGAGAT ATTCGGATG ATAATCTACC TCACCTTCCC TGTGGCTATG	60
TTCTGGGTTT CCAATCAGGC CGAGTGGTTT GAGGACGATG TCATACAGCG CAAGAGGGAG	120
CTGTGGCCAC CTGAGAAGCT TCAAGAGATA GAGGAATTCA AAGAGAGGTT ACGGAAGCGG	180
CGGGAGGAGA AGCTCCTTCG CGACGCCAG CAGAACTCC	219

【0086】

配列番号：16

配列の長さ：3516

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：リンホーマ

セルライン：U937

クローン名：HP10269

配列

ATGAGACCAT	TCTTCCTCTT	GTGTTTGCC	CTGCCTGGCC	TCCTGCATGC	CCAACAAGCC	60
TGCTCCCGTG	GGGCCTGCTA	TCCACCTGTT	GGGGACCTGC	TTGTTGGGAG	GACCCGGTT	120
CTCCGAGCTT	CATCTACCTG	TGGACTGACC	AAGCCTGAGA	CCTACTGCAC	CCAGTATGGC	180
GAGTGGCAGA	TGAAATGCTG	CAAGTGTGAC	TCCAGGCAGC	CTCACAACTA	CTACAGTCAC	240
CGAGTAGAGA	ATGTGGCTTC	ATCCTCCGGC	CCCATGCGCT	GGTGGCAGTC	CCAGAATGAT	300
GTGAACCCTG	TCTCTCTGCA	GCTGGACCTG	GACAGGAGAT	TCCAGCTTCA	AGAAGTCATG	360
ATGGAGTTCC	AGGGGCCAT	GCCTGCCGGC	ATGCTGATTG	AGCGCTCCTC	AGACTTCGGT	420
AAGACCTGGC	GAGTGTACCA	GTACCTGGCT	GCGACTGCA	CCTCCACCTT	CCCTCGGGTC	480
CGCCAGGGTC	GGCCTCAGAG	CTGGCAGGAT	GTTCGGTGCC	AGTCCCTGCC	TCAGAGGCCT	540
AATGCACGCC	TAAATGGGGG	GAAGGTCCAA	CTTAACCTTA	TGGATTTAGT	GTCTGGGATT	600
CCAGCAACTC	AAAGTAAAAA	AATTCAAGAG	GTGGGGGAGA	TCACAAACTT	GAGAGTCAAT	660
TTCACCAGGC	TGGCCCCTGT	GCCCCAAAGG	GGCTACCACC	CTCCCAGCGC	CTACTATGCT	720
GTGTCCCAGC	TCCGTCTGCA	GGGGAGCTGC	TTCTGTACG	GCCATGCTGA	TCGCTGCGCA	780
CCCCACCCCTC	CCCCCTCTCC	ACCCCTCTCC	ACCCCTCTCC	ACCTCCACCA	TCTCTCTCTC	840
CCCTGGAGAC	CGGGGGAGGG	CCAGGAUGGC	CATGAATGCC	AAAGGTGUGA	CTGUAATGGG	960
CACTCAGAGA	CATGTCACTT	TGACCCCGCT	GTGTTGCCG	CCAGCCAGGG	GGCATATGGA	1020

GGTGTGTGTG ACAATTGCCG GGACCACACC GAAGGCAAGA ACTGTGAGCG GTGTCAGCTG 1080
 CACTATTCC GGAACCGGCG CCCGGGAGCT TCCATTCAAG AGACCTGCAT CTCCTGCGAG 1140
 TGTGATCCGG ATGGGGCAGT GCCAGGGCT CCCTGTGACC CAGTGACCGG GCAGTGTGTG 1200
 TGCAAGGAGC ATGTGCAGGG AGAGCGCTGT GACCTATGCA AGCCGGGCTT CACTGGACTC 1260
 ACCTACGCCA ACCCGCAGGG CTGCCACCGC TGTGACTGCA ACATCCTGGG GTCCCGGAGG 1320
 GACATGCCGT GTGACGAGGA GAGTGGGCGC TGCCTTGTC TGCCCAACGT GGTGGGTCCC 1380
 AAATGTGACC AGTGTGCTCC CTACCACTGG AAGCTGGCCA GTGCCAGGG CTGTGAACCG 1440
 TGTGCCTGCG ACCCGCACAA CTCCCTCAGC CCACAGTGCA ACCAGTTCAC AGGGCAGTGC 1500
 CCCTGTGGG AAGGCTTG TGCCCTGATG TGCAGCGCTG CAGCCATCCG CCAGTGTCCA 1560
 GACCGGACCT ATGGAGACGT GGCCACAGGA TGCCGAGCCT GTGACTGTGA TTTCCGGGGA 1620
 ACAGAGGGCC CGGGCTGCGA CAAGGCATCA GGCGCTGCC TCTGCCGCC TGGCTTGACC 1680
 GGGCCCCGCT GTGACCAGTG CCAGCGAGGC TACTGCAATC GCTACCCGGT GTGCGTGGCC 1740
 TGCCACCCCTT GCTTCCAGAC CTATGATGCG GACCTCCGGG AGCAGGCCCT GCGCTTTGGT 1800
 AGACTCCGCA ATGCCACCGC CAGCCTGTGG TCAGGGCCTG GGCTGGAGGA CCGTGGCCTG 1860
 GCCTCCCGGA TCCTAGATGC AAAGAGTAAG ATTGAGCAGA TCCGAGCAGT TCTCAGCAGC 1920
 CCCGCAGTCA CAGAGCAGGA GGTGGCTCAG GTGGCCAGTG CCATCCTCTC CCTCAGGCCA 1980
 ACTCTCCAGG GCCTGCAGCT GGATCTGCC CTGGAGGAGG AGACGTTGTC CCTTCCGAGA 2040
 GACCTGGAGA GTCTTGACAG AAGCTTCAAT GGTCTCCTTA CTATGTATCA GAGGAAGAGG 2100
 GAGCAGTTG AAAAAATAAG CAGTGCTGAT CCTTCAGGAG CCTTCCGGAT GCTGAGCACA 2160
 GCCTACGAGC AGTCAGCCA GGCTGCTCAG CAGGTCTCCG ACAGCTCGCG CCTTTGGAC 2220
 CAGCTCAGGG ACAGCCGGAG AGAGGCAGAG AGGCTGGTGC GCCAGGCCGG AGGAGGAGGA 2280
 GGCACCGGCA GCCCAAGCT TGTGGCCCTG AGGCTGGAGA TGTCTTCGTT GCCTGACCTG 2340
 ACACCCACCT TCAACAAAGCT CTGTGGCAAC TCCAGGCAGA TGGCTTGCAC CCCAATATCA 2400
 TGCCCTGGTG AGCTATGTCC CCAAGACAAT GCCACAGCCT GTGGCTCCCG CTGCAGGGGT 2460
 GTCCTTCCA GGGCCGGTGG GGCCTTCTTG ATGGCGGGGC AGGTGGCTGA GCAGCTGCGG 2520
 GGCTTCAATG CCCAGCTCCA GCGGACCAGG CAGATGATTA GGGCAGCCGA GGAATCTGCC 2580
 TCACAGATTG AATCCAGTGC CCAGCGCTTG GAGACCCAGG TGAGCCAG CCGCTCCAG 2640
 ATGGAGGAAG ATGTCAAGACG CACACGGCTC CTAATCCAGC AGGTCCGGGA CTTCCTAAC 2700
 GACCCCGACA CTGATGCAGC CACTATCCAG GAGGTAGCG AGGCCGTGCT GGCCCTGTGG 2760

CTGCCACAG ACTCAGCTAC TGTTCTGCAG AAGATGAATG AGATCCAGGC CATTGCAGCC	2820
AGGCTCCCCA ACGTGGACTT GGTGCTGTCC CAGACCAAGC AGGACATTGC GCGTGCCCCGC	2880
CGGTTGCAGG CTGAGGCTGA GGAAGCCAGG AGCCGAGCCC ATGCAGTGGA GGGCCAGGTG	2940
GAAGATGTGG TTGGAACCT GCGGCAGGGG ACAGTGGCAC TGCAGGAAGC TCAGGACACC	3000
ATGCAAGGCA CCAGCCGCTC CCTTCGGCTT ATCCAGGACA GGGTTGCTGA GGTCAGCAG	3060
GTACTGCGGC CAGCAGAAAA GCTGGTGACA AGCATGACCA AGCAGCTGGG TGACTTCTGG	3120
ACACGGATGG AGGAGCTCCG CCACCAAGCC CGGCAGCAGG GGGCAGAGGC AGTCCAGGCC	3180
CAGCAGCTTG CGGAAGGTGC CAGCGAGCAG GCATTGAGTG CCCAAGAGGG ATTTGAGAGA	3240
ATAAAACAAA AGTATGCTGA GTTGAAGGAC CGGTTGGTC AGAGTTCCAT GCTGGGTGAG	3300
CAGGGTCCCC GGATCCACAG TCTGAAGACA GAGGCAGAGG AGCTGTTGG GGAGACCATG	3360
GAGATGATGG ACAGGATGAA AGACATGGAG TTGGAGCTGC TGCAGGGCAG CCAGGCCATC	3420
ATGCTGCGCT CAGCCGACCT GACAGGACTG GAGAAGCGTG TGGAGCAGAT CCGTGACCAC	3480
ATCAATGGGC GCGTGCTCTA CTATGCCACC TGCAAG	3516

【0087】

配列番号： 17

配列の長さ： 366

配列の型： 核酸

鎖の数： 二本鎖

トポロジー： 直鎖状

配列の種類： cDNA to mRNA

起源：

生物名： ホモ＝サピエンス

細胞の種類： 胃癌

クローン名： HP10298

配列

ATGGGCCTGT	TGCTCCTGGT	CCCATTGCTC	CTGCTGCCCG	GCTCCTACGG	ACTGCCCTTC	60
TACAACGGCT	TCTACTACTC	CAACAGCGCC	AACGACCAGA	ACCTAGGCAA	CGGTCATGGC	120
AAAGACCTCC	TTAATGGAGT	GAAGCTGGTG	GTGGAGACAC	CCGAGGGAGAC	CCTGTTCAACC	180
CGCATCCTAA	CTGTGGGCC	CCAGAGCCTG	GGGTCCGAAG	CTTTGGCTTC	CCCGACCCGC	240
AGAGCCGCTT	GTACGGTGT	TACTGCTACC	GCCAGCACTA	GGACCTGGGG	CCCTCCCCTG	300
CCGCATTCCC	TCACTGGCTG	TGTATTTATT	GAGTGGTTCG	TTTTCCCTTG	TGGGTTGGAG	360
CCATTT						366

【0088】

配列番号： 18

配列の長さ： 525

配列の型： 核酸

鎖の数： 二本鎖

トポロジー： 直鎖状

配列の種類： cDNA to mRNA

起源：

生物名： ホモ＝サピエンス

細胞の種類： 胃癌

クローン名： HP10368

配列

ATGGAGAAAA TTCCAGTGTC AGCATTCTTG CTCCTTGTGG CCCTCTCCTA CACTCTGGCC	60
AGAGATACCA CAGTCAAACC TGGAGCCAAA AAGGACACAA AGGACTCTCG ACCCAAAC TG	120
CCCCAGACCC TCTCCAGAGG TTGGGGTGAC CAACTCATCT GGACTCAGAC ATATGAAGAA	180
GCTCTATATA AATCCAAGAC AAGCAACAAA CCCTTGATGA TTATTCACTCA CTTGGATGAG	240
TGCCCACACA GTCAAGCTTT AAAGAAAGTG TTTGCTGAAA ATAAAGAAAT CCAGAAATTG	300
GCAGAGCAGT TTGTCCCTCCT CAATCTGGTT TATGAAACAA CTGACAAACA CCTTTCTCCT	360
GATGGCCAGT ATGTCCCCAG GATTATGTTT GTTGACCCAT CTCTGACAGT TAGAGCCGAT	420
ATCACTGGAA GATATTCAAA CCGTCTCTAT GCTTACGAAC CTGCAGATAAC AGCTCTGTTG	480
CTTGACAAACA TGAAGAAAGC TCTCAAGTTG CTGAAGACTG AATTG	525

【0089】

配列番号：19

配列の長さ：1296

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：纖維肉腫

セルライン：HT-1080

クローン名：HP00658

配列の特徴：

特徴を表す記号：CDS

存在位置：56..520

特徴を決定した方法：E

配列

CCTGCAGAGG ATCAAGACAG CACGTGGACC TCGCACAGCC TCTCCACAG GTACC ATG 58

Met

1

AAG GTC TCC GCG GCA GCC CTC GCT GTC ATC CTC ATT GCT ACT GCC CTC 106

Lys Val Ser Ala Ala Ala Leu Ala Val Ile Leu Ile Ala Thr Ala Leu

5

10

15

TGC GCT CCT GCA TCT GCC TCC CCA TAT TCC TCG GAC ACC ACA CCC TGC 154

Cys Ala Pro Ala Ser Ala Ser Pro Tyr Ser Ser Asp Thr Thr Pro Cys

20

25

30

TGC TTT GCC TAC ATT GCC CGC CCA CTG CCC CGT GCC CAC ATC AAG GAG 202

Cys Phe Ala Tyr Ile Ala Arg Pro Leu Pro Arg Ala His Ile Lys Glu

35

40

45

TAT TTC TAC ACC AGT GGC AAG TGC TCC AAC CCA GCA GTC GTC CAC AGG 250

Tyr Phe Tyr Thr Ser Gly Lys Cys Ser Asn Pro Ala Val Val His Arg
 50 55 60 65
 TCA AGG ATG CCA AAG AGA GAG GGA CAG CAA GTC TGG CAG GAT TTC CTG 298
 Ser Arg Met Pro Lys Arg Glu Gly Gln Gln Val Trp Gln Asp Phe Leu
 70 75 80
 TAT GAC TCC CGG CTG AAC AAG GGC AAG CTT TGT CAC CCG AAA GAA CCG 346
 Tyr Asp Ser Arg Leu Asn Lys Gly Lys Leu Cys His Pro Lys Glu Pro
 85 90 95
 CCA AGT GTG TGC CAA CCC AGA GAA GAA ATG GGT TCG GGA GTA CAT CAA 394
 Pro Ser Val Cys Gln Pro Arg Glu Glu Met Gly Ser Gly Val His Gln
 100 105 110
 CTC TTT GGA GAT GAG CTA GGA TGG AGA GTC CTT GAA CCT GAA CTT ACA 442
 Leu Phe Gly Asp Glu Leu Gly Trp Arg Val Leu Glu Pro Glu Leu Thr
 115 120 125
 CAA ATT TGC CTG TTT CTG CTT GCT CTT GTC CTA GCT TGG GAG GCT TCC 490
 Gln Ile Cys Leu Phe Leu Leu Ala Leu Val Leu Ala Trp Glu Ala Ser
 130 135 140 145
 CCT CAC TAT CCT ACC CCA CCC GCT CCT TGAAGGGCCC AGA 530
 Pro His Tyr Pro Thr Pro Pro Ala Pro
 150
 TTCTACCACA CAGCAGCAGT TACAAAAACC TTCCCCAGGC TGGACGTGGT GGCTCACGCC 590
 TGTAATCCA GCACTTGGG AGGCCAAGGT GGGTGGATCA CTTGAGGTCA GGAGTTCGAG 650
 ACCAGCCTGG CCAACATGAT GAAACCCAT CTCTACTAAA AATACAAAAA ATTAGCCGGG 710
 CGTGGTAGCG GGCGCCTGTA GTCCCAGCTA CTCGGGAGGC TGAGGCAGGA GAATGGCGTG 770
 AACCCGGGAG GCGGAGCTTG CAGTGAGCCG AGATCGCGCC ACTGCACTCC AGCCTGGCG 830
 AACACACAC AACACACAC AACACACAC AACACACAC AACACACAC ATTACCCCCC 890

 AACCCAGGAG GTGGAGGCTG CAGTGAGCTG AGATTGTGCC ACTTCACGCC AGCCTGGGTG 910
 ACAAAAGTGAG ACTCCGTCAC AACAAACAACA ACAAAAAAGCT TCCCCAACTA AAGCCTAGAA 1070

特平 8-243060

GAGCTTCTGA GGCGCTGCTT TGTCAAAAGG AAGTCTCTAG GTTCTGAGCT CTGGCTTGC 1130
CTTGGCTTG CCAGGGCTCT GTGACCAGGA AGGAAGTCAG CATGCCTCTA GAGGCAAGGA 1190
GGGGAGGAAC GCTGCACTCT TAAGCTTCCG CCGTCTAAC CCCTCACAGG AGCTTACTGG 1250
CAAACATGAA AAATCGGCTT ACCATTAAAG TTCTCAATGC AACCAT 1296

【0090】

配列番号：20

配列の長さ：3311

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：類表皮癌

セルライン：KB

クローン名：HP00714

配列の特徴：

特徴を表す記号：CDS

存在位置：57..1004

特徴を決定した方法：E

配列

GAGCGGCGGC CACGGCATCC TGTGCTGTGG GGGCTACGAG GAAAGATCTA ATTATC ATG 59

Met

1

GAC CTG CGA CAG TTT CTT ATG TGC CTG TCC CTG TGC ACA GCC TTT GCC 107

Asp Leu Arg Gln Phe Leu Met Cys Leu Ser Leu Cys Thr Ala Phe Ala

5

10

15

TTG AGC AAA CCC ACA GAA AAG AAG GAC CGT GTA CAT CAT GAG CCT CAG 155

Leu Ser Lys Pro Thr Glu Lys Lys Asp Arg Val His His Glu Pro Gln

20

25

30

CTG ACT CAC AAC CTT GAC ATT CAT CCT GAC ACT TTT CAT TAT CAC CAT 200

55

40

45

GAT GCC TTC TTG GGT GCT GAA GAA GCA AAG ACC TTT GAT CAG CTG ACA 251

Asp Ala Phe Leu Gly Ala Glu Glu Ala Lys Thr Phe Asp Gln Leu Thr			
50	55	60	65
CCA GAA GAG AGC AAG GAA AGG CTT GGA AAG ATT GTA AGT AAA ATA GAT			299
Pro Glu Glu Ser Lys Glu Arg Leu Gly Lys Ile Val Ser Lys Ile Asp			
70	75	80	
GGC GAC AAG GAC GGG TTT GTC ACT GTG GAT GAG CTC AAA GAC TGG ATT			347
Gly Asp Lys Asp Gly Phe Val Thr Val Asp Glu Leu Lys Asp Trp Ile			
85	90	95	
AAA TTT GCA CAA AAG CGC TGG ATT TAC GAG GAT GTA GAG CGA CAG TGG			395
Lys Phe Ala Gln Lys Arg Trp Ile Tyr Glu Asp Val Glu Arg Gln Trp			
100	105	110	
AAG GGG CAT GAC CTC AAT GAG GAC GGC CTC GTT TCC TGG GAG GAG TAT			443
Lys Gly His Asp Leu Asn Glu Asp Gly Leu Val Ser Trp Glu Glu Tyr			
115	120	125	
AAA AAT GCC ACC TAC GGC TAC GTT TTA GAT GAT CCA GAT CCT GAT GAT			491
Lys Asn Ala Thr Tyr Gly Tyr Val Leu Asp Asp Pro Asp Pro Asp Asp			
130	135	140	145
GGA TTT AAC TAT AAA CAG ATG ATG GTT AGA GAT GAG CGG AGG TTT AAA			539
Gly Phe Asn Tyr Lys Gln Met Met Val Arg Asp Glu Arg Arg Phe Lys			
150	155	160	
ATG GCA GAC AAG GAT GGA GAC CTC ATT GCC ACC AAG GAG GAG TTC ACA			587
Met Ala Asp Lys Asp Gly Asp Leu Ile Ala Thr Lys Glu Glu Phe Thr			
165	170	175	
GCT TTC CTG CAC CCT GAG GAG TAT GAC TAC ATG AAA GAT ATA GTA GTA			635
Ala Phe Leu His Pro Glu Glu Tyr Asp Tyr Met Lys Asp Ile Val Val			
180	185	190	
CAG GAA ACA ATG GAA GAT ATA GAT AAG AAT GCT GAT GGT TTC ATT GAT			683
Gln Glu Thr Met Glu Asp Ile Asp Lys Asn Ala Asp Gly Phe Ile Asp			
195	200	205	

AAGTCCAACA	TTTCATAGGT	AGTAGGGGCC	ACATATTACA	TTCAGTTGCT	ATAGGTCCAG	1560
CAACTGAACC	TGCCATTACC	TGGGCAAGGA	AAGATCCCTT	TGCTCTAGGA	AAGCTTGGCC	1620
CAAATTGATT	TTCTTCTTT	TCCCCCTGTA	GGACTGACTG	TTGGCTAATT	TTGTCAAGCA	1680
CAGCTGTGGT	GGGAAGAGTT	AGGGCCAGTG	TCTTGAAAAT	CAATCAAGTA	GTGAATGTGA	1740
TCTCTTGCA	GAGCTATAGA	TAGAAACAGC	TGGAAAAC	AAGGAAAAAT	ACAAGTGT	1800
TCGGGGCATA	CATTTTTT	CTGGGTGTGC	ATCTGTTGAA	ATGCTCAAGA	CTTAATTATT	1860
TGCCTTTGA	AATCACTGTA	AATGCCCCA	TCCGGTTCC	CTTCTTCCC	GGTGTGCCAA	1920
GGAATTAATC	TTGGTTTCAC	TACAATTAAA	ATTCACTCCT	TTCCAATCAT	GTCATTGAAA	1980
GTGCCTTAA	CGAAAGAAAT	GGTCACTGAA	TGGGAATTCT	CTTAAGAAC	CCTGAGATTA	2040
AAAAAAAGACT	ATTTGGATAA	CTTATAGGAA	AGCCTAGAAC	CTCCCAGTAG	AGTGGGGATT	2100
TTTTCTTCT	TCCCTTCTC	TTTGGACAA	TAGTTAAATT	AGCAGTATTA	GTTATGAGTT	2160
TGGTTGCAGT	GTTCTTATCT	TGTGGGCTGA	TTTCCAAAAAA	CCACATGCTG	CTGAATTAC	2220
CAGGGATCCT	CATACCTCAC	AATGCAAACC	ACTTACTACC	AGGCCTTTT	CTGTGTCCAC	2280
TGGAGAGCTT	GAGCTCACAC	TCAAAGATCA	GAGGACCTAC	AGAGAGGGCT	CTTGGTTG	2340
AGGACCATGG	CTTACCTTTC	CTGCCTTGA	CCCATCACAC	CCCATTCT	CCTCTTCCC	2400
TCTCCCCGCT	GCCAAAAAAA	AAAAAAAAG	GAAACGTTA	TCATGAATCA	ACAGGGTTTC	2460
AGTCCTTATC	AAAGAGAGAT	GTGGAAAGAG	CTAAAGAAC	CACCCITTGT	TCCCAACTCC	2520
ACTTTACCCA	TATTTATGC	AACACAAACA	CTGTCCTT	GGGTCCCTT	CTTACAGATG	2580
GACCTCTTGA	GAAGAATTAT	CGTATTCCAC	GTTTTAGCC	CTCAGGTTAC	CAAGATAAAAT	2640
ATATGTATAT	ATAACCTTA	TTATTGCTAT	ATCTTGTGG	ATAATACATT	CAGGTGGTGC	2700
TGGGTGATT	ATTATAATCT	GAACCTAGGT	ATATCCTTG	GTCTCCACA	GTCATGTTGA	2760
GGTGGGCTCC	CTGGTATGGT	AAAAAGCCAG	GTATAATGTA	ACTTCACCCC	AGCCTTGTA	2820
CTAAGCTCTT	GATAGTGGAT	ATACTCTT	AAGTTAGCC	CCAATATAGG	GTAATGGAAA	2880
TTCCCTGCC	TCTGGGTCC	CCATTTCAC	TATTAAGAAG	ACCAAGATA	ATTTAATAAT	2940
GCCACCAACT	CTGGCTTAGT	TAAGTGAGAG	TGTGAACGT	GTGGCAAGAG	AGCCTCACAC	3000
CTCACTAGGT	GCAGAGAGCC	CAGGCCTTAT	GTTAAAATCA	TGCACCTGAA	AAGCAAACCT	3060
TAATCTGCAA	AGACAGCAGC	AAGCATTATA	CGGTACATCTT	GAATGATCCC	TTTGAAATT	3120
TTTTTTGTT	TGTTTGT	AATCAAGCCT	GAGGCTGGTG	AACAGTAGCT	ACACACCCAT	3180
ATTGTGTGTT	CTGTGAATGC	TAGCTTCTT	GAATTGGAT	ATTGGTTATT	TTTATAGAG	3240

特平 8-243060

TGTAAACCAA GTTTATATT CTGCAATGCG AACAGGTACC TATCTGTTTC TAAATAAAAC 3300

TGTTTACATT C 3311

【0091】

配列番号：21

配列の長さ：1152

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：HP00876

配列の特徴：

特徴を表す記号：CDS

存在位置：147..623

特徴を決定した方法：E

配列

ACTGGAGACA	CTGAAGAAGG	CAGGGGCCCT	TAGAGTCTTG	GTTGCCAAAC	AGATTGCAG	60
ATCAAGGAGA	ACCCAGGAGT	TTCAAAGAAG	CGCTAGTAAG	GTCTCTGAGA	TCCTTGCACT	120
AGCTACATCC	TCAGGGTAGG	AGGAAG	ATG GCT TCC	AGA AGC ATG CGG CTG CTC	173	
Met Ala Ser Arg Ser Met Arg Leu Leu						
	1		5			
CTA TTG CTG AGC TGC CTG GCC	AAA ACA GGA GTC	CTG GGT GAT ATC ATC				221
Leu Leu Leu Ser Cys Leu Ala Lys Thr Gly Val Leu Gly Asp Ile Ile						
10	15	20		25		
ATG AGA CCC AGC TGT GCT CCT GGA TGG TTT TAC CAC AAG TCC AAT TGC						269
Met Arg Pro Ser Cys Ala Pro Gly Trp Phe Tyr His Lys Ser Asn Cys						
30	35		40			
TAT GGT TAC TTC AGG AAG CTG AGG AAC TGG TCT GAT GCC GAG CTC GAG						317
Tyr Gly Tyr Phe Arg Lys Leu Arg Asn Trp Ser Asp Ala Glu Leu Glu						
45	50		55			

TGT CAG TCT TAC GGA AAC GGA GCC CAC CTG GCA TCT ATC CTG AGT TTA 365
 Cys Gln Ser Tyr Gly Asn Gly Ala His Leu Ala Ser Ile Leu Ser Leu
 60 65 70
 AAG GAA GCC AGC ACC ATA GCA GAG TAC ATA AGT GGC TAT CAG AGA AGC 413
 Lys Glu Ala Ser Thr Ile Ala Glu Tyr Ile Ser Gly Tyr Gln Arg Ser
 75 80 85
 CAG CCG ATA TGG ATT GGC CTG CAC GAC CCA CAG AAG AGG CAG CAG TGG 461
 Gln Pro Ile Trp Ile Gly Leu His Asp Pro Gln Lys Arg Gln Gln Trp
 90 95 100 105
 CAG TGG ATT GAT GGG GCC ATG TAT CTG TAC AGA TCC TGG TCT GGC AAG 509
 Gln Trp Ile Asp Gly Ala Met Tyr Leu Tyr Arg Ser Trp Ser Gly Lys
 110 115 120
 TCC ATG GGT GGG AAC AAG CAC TGT GCT GAG ATG AGC TCC AAT AAC AAC 557
 Ser Met Gly Gly Asn Lys His Cys Ala Glu Met Ser Ser Asn Asn Asn
 125 130 135
 TTT TTA ACT TGG AGC AGC AAC GAA TGC AAC AAG CGC CAA CAC TTC CTG 605
 Phe Leu Thr Trp Ser Ser Asn Glu Cys Asn Lys Arg Gln His Phe Leu
 140 145 150
 TGC AAG TAC CGA CCA TAGAGCAAGA ATCAAGATTG TGCTAACTCC 650
 Cys Lys Tyr Arg Pro
 155
 TGCACAGCCC CGTCCTCTTC CTTTCTGCTA GCCTGGCTAA ATCTGCTCAT TATTTCAGAG 710
 GGGAAACCTA GCAAACTAAG AGTGATAAGG GCCCTACTAC ACTGGCTTT TTAGGCTTAG 770
 AGACAGAAAC TTTAGCATTG GCCCAGTAGT GGCTTCTAGC TCTAAATGTT TGCCCCGCCA 830
 TCCCTTTCCA CAGTATCCTT CTTCCCTCCT CCCCTGTCTC TGGCTGTCTC GAGCAGTCTA 890
 CAAACACTCCA TCTCCAGCCT ATCAAACACCC TCCCTCTTTTC CCCATAACAA CTAAACATT 950
 CAGGCGCTT TCTATGGCT GCAAGGCAACG CAGGCGACAC AACCTCTGCT TGTTTTTCCT 1070
 TTGGCCATGG GAAGGTTAC CAGTAGAATC CTTGCTAGGT TGATGTGGGC CATAACATTCC 1130

特平 8-243060

TTTAATAAAC CATTGTGTAC AT

1152

【0092】

配列番号： 22

配列の長さ： 1749

配列の型： 核酸

鎖の数： 二本鎖

トポロジー： 直鎖状

配列の種類： cDNA to mRNA

起源：

生物名： ホモ＝サピエンス

細胞の種類： 肝臓

クローン名： HPO1134

配列の特徴：

特徴を表す記号： CDS

存在位置： 117..1247

特徴を決定した方法： E

配列

AATCACAGCA	GTNCCGACGT	CGTGGGTGTT	TGGTGTGAGG	CTGCGAGCCG	CCGCCGCCAC	60
CACTGCCACC	ACGGTCGCCT	GCCACAGGTG	TCTGCAATTG	AACTCCAAGG	TGCAGA ATG	119

Met

1

GTT TGG AAA GTA GCT GTA TTC CTC AGT GTG GCC CTG GGC ATT GGT GCC	167
Val Trp Lys Val Ala Val Phe Leu Ser Val Ala Leu Gly Ile Gly Ala	

5

10

15

GTT CCT ATA GAT GAT CCT GAA GAT GGA GGC AAG CAC TGG GTG GTG ATC	215
Val Pro Ile Asp Asp Pro Glu Asp Gly Gly Lys His Trp Val Val Ile	

20

25

30

CTG CCA GGT TCA AAT GGC TGC TAT AAT TAT ACC CAC CAC CCA CAC CCC	262
---	-----

TGC CAT GCC TAC CAG ATC ATT CAC CGC AAT GGG ATT CCT GAC GAA CAG	311
---	-----

Cys	His	Ala	Tyr	Gln	Ile	Ile	His	Arg	Asn	Gly	Ile	Pro	Asp	Glu	Gln	
50		55					60					65				
ATC	GTT	GTG	ATG	ATG	TAC	GAT	GAC	ATT	GCT	TAC	TCT	GAA	GAC	AAT	CCC	359
Ile	Val	Val	Met	Met	Tyr	Asp	Asp	Ile	Ala	Tyr	Ser	Glu	Asp	Asn	Pro	
								70		75			80			
ACT	CCA	GGA	ATT	GTG	ATC	AAC	AGG	CCC	AAT	GGC	ACA	GAT	GTC	TAT	CAG	407
Thr	Pro	Gly	Ile	Val	Ile	Asn	Arg	Pro	Asn	Gly	Thr	Asp	Val	Tyr	Gln	
								85		90			95			
GGA	GTC	CCG	AAG	GAC	TAC	ACT	GGA	GAG	GAT	GTT	ACC	CCA	CAA	AAT	TTC	455
Gly	Val	Pro	Lys	Asp	Tyr	Thr	Gly	Glu	Asp	Val	Thr	Pro	Gln	Asn	Phe	
								100		105			110			
CTT	GCT	GTG	TTG	AGA	GGC	GAT	GCA	GAA	GCA	GTG	AAG	GGC	ATA	GGA	TCC	503
Leu	Ala	Val	Leu	Arg	Gly	Asp	Ala	Glu	Ala	Val	Lys	Gly	Ile	Gly	Ser	
								115		120			125			
GGC	AAA	GTC	CTG	AAG	AGT	GGC	CCC	CAG	GAT	CAC	GTG	TTC	ATT	TAC	TTC	551
Gly	Lys	Val	Leu	Lys	Ser	Gly	Pro	Gln	Asp	His	Val	Phe	Ile	Tyr	Phe	
								130		135			140			145
ACT	GAC	CAT	GGA	TCT	ACT	GGA	ATA	CTG	GTT	TTT	CCC	AAT	GAA	GAT	CTT	599
Thr	Asp	His	Gly	Ser	Thr	Gly	Ile	Leu	Val	Phe	Pro	Asn	Glu	Asp	Leu	
								150		155			160			
CAT	GTA	AAG	GAC	CTG	AAT	GAG	ACC	ATC	CAT	TAC	ATG	TAC	AAA	CAC	AAA	647
His	Val	Lys	Asp	Leu	Asn	Glu	Thr	Ile	His	Tyr	Met	Tyr	Lys	His	Lys	
								165		170			175			
ATG	TAC	CGA	AAG	ATG	GTG	TTC	TAC	ATT	GAA	GCC	TGT	GAG	TCT	GGG	TCC	695
Met	Tyr	Arg	Lys	Met	Val	Phe	Tyr	Ile	Glu	Ala	Cys	Glu	Ser	Gly	Ser	
								180		185			190			
ATG	ATG	AAC	CAC	CTG	CCG	GAT	AAC	ATC	AAT	GTT	TAT	GCA	ACT	ACT	GCT	743
Met	Met	Asn	His	Leu	Pro	Asp	Asn	Ile	Asn	Val	Tyr	Ala	Thr	Thr	Ala	
								195		200			205			

GCC AAC CCC AGA GAG TCG TCC TAC GCC TGT TAC TAT GAT GAG AAG AGG	791		
Ala Asn Pro Arg Glu Ser Ser Tyr Ala Cys Tyr Tyr Asp Glu Lys Arg			
210	215	220	225
TCC ACG TAC CTG GGG GAC TGG TAC AGC GTC AAC TGG ATG GAA GAC TCG	839		
Ser Thr Tyr Leu Gly Asp Trp Tyr Ser Val Asn Trp Met Glu Asp Ser			
230	235	240	
GAC GTG GAA GAT CTG ACT AAA GAG ACC CTG CAC AAG CAG TAC CAC CTG	887		
Asp Val Glu Asp Leu Thr Lys Glu Thr Leu His Lys Gln Tyr His Leu			
245	250	255	
GTA AAA TCG CAC ACC AAC ACC AGC CAC GTC ATG CAG TAT GGA AAC AAA	935		
Val Lys Ser His Thr Asn Thr Ser His Val Met Gln Tyr Gly Asn Lys			
260	265	270	
ACA ATC TCC ACC ATG AAA GTG ATG CAG TTT CAG GGT ATG AAA CGC AAA	983		

Thr Ile Ser Thr Met Lys Val Met Gln Phe Gln Gly Met Lys Arg Lys
 275 280 285
 GCC AGT TCT CCC GTC CCC CTA CCT CCA GTC ACA CAC CTT GAC CTC ACC 1031
 Ala Ser Ser Pro Val Pro Leu Pro Val Thr His Leu Asp Leu Thr
 290 295 300 305
 CCC AGC CCT GAT GTG CCT CTC ACC ATC ATG AAA AGG AAA CTG ATG AAC 1079
 Pro Ser Pro Asp Val Pro Leu Thr Ile Met Lys Arg Lys Leu Met Asn
 310 315 320
 ACC AAT GAT CTG GAG GAG TCC AGG CAG CTC ACG GAG GAG ATC CAG CGG 1127
 Thr Asn Asp Leu Glu Glu Ser Arg Gln Leu Thr Glu Glu Ile Gln Arg
 325 330 335
 CAT CTG GAT TAC GAG TAT GCG TTG AGA CAT TTG TAC GTG CTG GTC AAC 1175
 His Leu Asp Tyr Glu Tyr Ala Leu Arg His Leu Tyr Val Leu Val Asn
 340 345 350
 CTT TGT GAG AAG CCG TAT CCG CTT CAC AGG ATA AAA TTG TCC ATG GAC 1223
 Leu Cys Glu Lys Pro Tyr Pro Leu His Arg Ile Lys Leu Ser Met Asp
 355 360 365
 CAC GTG TGC CTT GGT CAC TAC TGAAGAGCTG CCTCCTGGAA GCTTTT 1270
 His Val Cys Leu Gly His Tyr
 370 375
 CCAAGTGTGA GCGCCCCACC GACTGTGTGC TGATCAGAGA CTGGAGAGGT GGAGTGAGAA 1330
 GTCTCCGCTG CTCGGGCCCT CCTGGGGAGC CCCCCGCTCCA GGGCTCGCTC CAGGACCTTC 1390
 TTCACAAAGAT GACTTGCTCG CTGTTACCTG CTTCCCCAGT CTTTCTGAA AAACATACAAA 1450
 TTAGGGTGGG AAAAGCTCTG TATTGAGAAG GGTCAATATT GCTTTCTAGG AGGTTTGTG 1510
 TTTGCCTGT TAGTTTGAG GAGCAGGAAG CTCATGGGGG CTTCTGTAGC CCCTCTCAAA 1570
 AGGAGTCTTT ATTCTGAGAA TTTGAAGCTG AAACCTCTTT AAATCTTCAG AATGATTTA 1630
 TTGAAGAGGG CCGCAAGCCC CAAATGGAAA ACTGTTTTA GAAAATATGA TGATTTTGA 1690
 TTGCTTTGT ATTTAATTCT GCAGGTGTTC AAGTCTAAA AAATAAAGAT TTATAACAG 1749

【0093】

配列番号: 23

配列の長さ: 988

配列の型: 核酸

鎖の数: 二本鎖

トポロジー: 直鎖状

配列の種類: cDNA to mRNA

起源:

生物名: ホモ=サピエンス

細胞の種類: 類表皮癌

セルライン: KB

クローン名: HP10029

配列の特徴:

特徴を表す記号: CDS

存在位置: 9..530

特徴を決定した方法: E

配列

AGTCCAAC ATG GCG GCG CCC AGC GGA GGG TGG AAC GGC GTC CGC GCG AGC 50

Met Ala Ala Pro Ser Gly Gly Trp Asn Gly Val Arg Ala Ser

1 5 10

TTG TGG GCC GCG CTG CTC CTA GGG GCC GTG GCG CTG AGG CCG GCG GAG 98

Leu Trp Ala Ala Leu Leu Leu Gly Ala Val Ala Leu Arg Pro Ala Glu

15 20 25 30

GCG GTG TCC GAG CCC ACG ACC GTG GCG TTT GAC GTG CGG CCC GGC GGC 146

Ala Val Ser Glu Pro Thr Thr Val Ala Phe Asp Val Arg Pro Gly Gly

35 40 45

ATG ATG CTT TCC TTT TGG CTC ACG CTC CCC CCC CCC CAC AAA TAT ACC 191

50 55 60

TGT ATG TTC ACT TAC GCC TCT CAA GGA GGG ACC AAT GAG CAA TGG CAG 242

Cys	Met	Phe	Thr	Tyr	Ala	Ser	Gln	Gly	Gly	Thr	Asn	Glu	Gln	Trp	Gln	
65																
ATG	AGT	CTG	GGG	ACC	AGC	GAA	GAC	CAC	CAG	CAC	TTC	ACC	TGC	ACC	ATC	290
Met	Ser	Leu	Gly	Thr	Ser	Glu	Asp	His	Gln	His	Phe	Thr	Cys	Thr	Ile	
80																
TGG	AGG	CCC	CAG	GGG	AAG	TCC	TAT	CTG	TAC	TTC	ACA	CAG	TTC	AAG	GCA	338
Trp	Arg	Pro	Gln	Gly	Lys	Ser	Tyr	Leu	Tyr	Phe	Thr	Gln	Phe	Lys	Ala	
95																
GAG	GTG	CGG	GGC	GCT	GAG	ATT	GAG	TAC	GCC	ATG	GCC	TAC	TCT	AAA	GCC	386
Glu	Val	Arg	Gly	Ala	Glu	Ile	Glu	Tyr	Ala	Met	Ala	Tyr	Ser	Lys	Ala	
115																
GCA	TTT	GAA	AGG	GAA	AGT	GAT	GTC	CCT	CTG	AAA	ACT	GAG	GAA	TTT	GAA	434
Ala	Phe	Glu	Arg	Glu	Ser	Asp	Val	Pro	Leu	Lys	Thr	Glu	Glu	Phe	Glu	
130																
GTG	ACC	AAA	ACA	GCA	GTG	GCT	CAC	AGG	CCC	GGG	GCA	TTC	AAA	GCT	GAG	482
Val	Thr	Lys	Thr	Ala	Val	Ala	His	Arg	Pro	Gly	Ala	Phe	Lys	Ala	Glu	
145																
CTG	TCC	AAG	CTG	GTG	ATT	GTG	GCC	AAG	GCA	TCG	CGC	ACT	GAG	CTG		527
Leu	Ser	Lys	Leu	Val	Ile	Val	Ala	Lys	Ala	Ser	Arg	Thr	Glu	Leu		
160																
TGA	CCAGCAGCCC	TGTTGCGGGT	GGCACCTTCT	CATCTCCGGT	GAAGCTGAAG											580
GGGCCTGTGG	CCCTGAAAGG	GCCAGCACAT	CACTGGTTTT	CTAGGAGGGA	CTCTTAAGTT											640
TTCTACCTGG	GCTGACGTTG	CCTTGTCCGG	AGGGGCTTGC	AGGGTGGCTG	AAGCCCTGGG											700
GCAGAGAACAA	GAGGCTCCAG	GGCCCTCCTG	GCTCCCAACA	GCTTCTCAGT	TCCCACCTCC											760
TGCTGAGCTC	TTCTGGACTC	AGGATCGCAG	ATCCGGGGCA	CAAAGAGGGT	GGGAAACATG											820
GGGGCTATGC	TGGGGAAAGC	AGCCATGCTC	CCCCCGACCT	CCAGCCGAGC	ATCCTTCATG											880
AGCCTGCAGA	ACTGCTTCC	TATGTTACC	CAGGGGACCT	CCTTCAGAT	GAACTGGGAA											940
GAGATGAAAT	GTTCAT	ATTTAAATAA	ATAAGAACAT	TAAAAAGC												988

【0094】

配列番号：24

配列の長さ：390

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：類表皮癌

セルライン：KB

クローン名：HP10189

配列の特徴：

特徴を表す記号：CDS

存在位置：102..323

特徴を決定した方法：E

配列

AATCAGCTTC AGCAATGGAG CGTGCAAAAC ACCAGTGAGC TTCTGTCTTG CTGGAGGGTC 60

GGCTTGCGGC GGAAGTGGCT TTGTTGACCG GGAGAACGA G ATG GGG GTG AAG CTG 116

Met Gly Val Lys Leu

1 5

GAG ATA TTT CGG ATG ATA ATC TAC CTC ACT TTC CCT GTG GCT ATG TTC 164

Glu Ile Phe Arg Met Ile Ile Tyr Leu Thr Phe Pro Val Ala Met Phe

10 15 20

TGG GTT TCC AAT CAG GCC GAG TGG TTT GAG GAC GAT GTC ATA CAG CGC 212

Trp Val Ser Asn Gln Ala Glu Trp Phe Glu Asp Asp Val Ile Gln Arg

25 30 35

Lys Arg Ala Leu Trp Pro Pro Glu Lys Leu Gln Glu Ile Glu Glu The

40 45 50

特平 8-243060

AAA GAG AGG TTA CGG AAG CGG CGG GAG AAG CTC CTT CGC GAC GCC 308
Lys Glu Arg Leu Arg Lys Arg Arg Glu Glu Lys Leu Leu Arg Asp Ala

55 60 65

CAG CAG AAC TCC TGAGGCCTCC AAGTGGGAGT CCTAGCCCCT 350
Gln Gln Asn Ser

70

CCCCTGATGA AATATACATA TACTCAGTTC CTTGTTATT 390
【0095】

配列番号：25

配列の長さ：4667

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：リンホーマ

セルライン：U937

クローン名：HP10269

配列の特徴：

特徴を表す記号：CDS

存在位置：754..4272

特徴を決定した方法：E

配列

CATTTAGTTA	CTCTGCTCAT	TTCTCTTAAG	CTTCCTTGG	ATGAGTTGAG	CTTGAAATCC	60
TTCCTGATGA	ACCTTGCCTT	TTAAGGATCC	TCCAAATGCC	CCAAGAAGCT	GGGATTTTC	120
ATTTTTTTT	TCACTGGGGA	GGGAATGGT	GCTTCCAGG	GTCCTGGATG	TTTGAGTCTT	180
CTCACCTTCC	AGCCCGGTGA	TATGTCTGGA	GCTTTAACTC	TCTATATAAG	CCCTAATCTT	240
TGTGTTCTCT	GCCTGATCTT	CTGTCTGGGG	TGGTCCAGGT	CACAAGAAGA	AGCTGACCCC	300
TGCTGGCTTT	GGGAAAATGC	TGAGTTCATT	GCCTGGCACA	AATGCAAGGG	CCCTTCCCCA	360
CCCTGTGAAT	TCTGGTCTCT	GATGATCACT	TACATGTGCC	TTGTGCTTTC	TGTTGAGGG	420
GCCCCTTGCA	GCCCCCACAG	GCAGGTGGGC	ATTGTGGAGC	TCACTACAAG	AACTCTGGGA	480
CCGACCGACC	AACCCACTTG	CCCAGTCCCG	TCCTGGGAGG	TGGGGTGCA	GTGACGACAG	540
CCCCCTCTCA	CCCCCTCCATC	TTTCTCACTA	CCCCCCCCCTCC	CCTCCCCCTA	CACCCACCCCA	600
CATTCAGAG	GAGCTTCAAG	GUGATCTGGA	GAAAGAACGG	CAGAACACAC	AGCAAGGAAA	720
GGTCCTTCT	GGGGATCACC	CCATTGGCTG	AAG ATG AGA CCA	TTC TTC CTC	TTG	774

			Met	Arg	Pro	Phe	Phe	Leu	Leu							
			1	5												
TGT	TTT	GCC	CTG	CCT	GGC	CTC	CTG	CAT	GCC	CAA	CAA	GCC	TGC	TCC	CGT	822
Cys	Phe	Ala	Leu	Pro	Gly	Leu	Leu	His	Ala	Gln	Gln	Ala	Cys	Ser	Arg	
10																
GGG	GCC	TGC	TAT	CCA	CCT	GTT	GGG	GAC	CTG	CTT	GTT	GGG	AGG	ACC	CGG	870
Gly	Ala	Cys	Tyr	Pro	Pro	Val	Gly	Asp	Leu	Leu	Val	Gly	Arg	Thr	Arg	
25																
TTT	CTC	CGA	GCT	TCA	TCT	ACC	TGT	GGA	CTG	ACC	AAG	CCT	GAG	ACC	TAC	918
Phe	Leu	Arg	Ala	Ser	Ser	Thr	Cys	Gly	Leu	Thr	Lys	Pro	Glu	Thr	Tyr	
40																
TGC	ACC	CAG	TAT	GGC	GAG	TGG	CAG	ATG	AAA	TGC	TGC	AAG	TGT	GAC	TCC	966
Cys	Thr	Gln	Tyr	Gly	Glu	Trp	Gln	Met	Lys	Cys	Cys	Lys	Cys	Asp	Ser	
55																
AGG	CAG	CCT	CAC	AAC	TAC	TAC	AGT	CAC	CGA	GTA	GAG	AAT	GTG	GCT	TCA	1014
Arg	Gln	Pro	His	Asn	Tyr	Tyr	Ser	His	Arg	Val	Glu	Asn	Val	Ala	Ser	
60																
AGG	CAG	CCT	CAC	AAC	TAC	TAC	AGT	CAC	CGA	GTA	GAG	AAT	GTG	GCT	TCA	1062
Ser	Ser	Gly	Pro	Met	Arg	Trp	Trp	Gln	Ser	Gln	Asn	Asp	Val	Asn	Pro	
75																
TCC	TCC	GGC	CCC	ATG	CGC	TGG	TGG	CAG	TCC	CAG	AAT	GAT	GTG	AAC	CCT	1110
Ser	Ser	Gly	Pro	Met	Arg	Trp	Trp	Gln	Ser	Gln	Asn	Asp	Val	Asn	Pro	
90																
GTC	TCT	CTG	CAG	CTG	GAC	CTG	GAC	AGG	AGA	TTC	CAG	CTT	CAA	GAA	GTC	1158
Val	Ser	Leu	Gln	Leu	Asp	Leu	Asp	Arg	Arg	Phe	Gln	Leu	Gln	Glu	Val	
105																
ATG	ATG	GAG	TTC	CAG	GGG	CCC	ATG	CCT	GCC	GGC	ATG	CTG	ATT	GAG	CGC	1206
Met	Met	Glu	Phe	Gln	Gly	Pro	Met	Pro	Ala	Gly	Met	Leu	Ile	Glu	Arg	
120																
TCC	TCA	GAC	TTC	GGT	AAG	ACC	TGG	CGA	GTG	TAC	CAG	TAC	CTG	GCT	GCC	140
Ser	Ser	Asp	Phe	Gly	Lys	Thr	Trp	Arg	Val	Tyr	Gln	Tyr	Leu	Ala	Ala	
135																
145																150

GAC TGC ACC TCC ACC TTC CCT CGG GTC CGC CAG GGT CGG CCT CAG AGC	1254
Asp Cys Thr Ser Thr Phe Pro Arg Val Arg Gln Gly Arg Pro Gln Ser	
155 160 165	
TGG CAG GAT GTT CGG TGC CAG TCC CTG CCT CAG AGG CCT AAT GCA CGC	1302
Trp Gln Asp Val Arg Cys Gln Ser Leu Pro Gln Arg Pro Asn Ala Arg	
170 175 180	
CTA AAT GGG GGG AAG GTC CAA CTT AAC CTT ATG GAT TTA GTG TCT GGG	1350
Leu Asn Gly Gly Lys Val Gln Leu Asn Leu Met Asp Leu Val Ser Gly	
185 190 195	
ATT CCA GCA ACT CAA ACT CAA AAA ATT CAA CAG GTG GGG GAG ATC ACA	1398
Ile Pro Ala Thr Gln Ser Gln Lys Ile Gln Glu Val Gly Glu Ile Thr	
200 205 210 215	
AAC TTG AGA GTC AAT TTC ACC AGG CTG GCC CCT GTG CCC CAA AGG GGC	1446
Asn Leu Arg Val Asn Phe Thr Arg Leu Ala Pro Val Pro Gln Arg Gly	
220 225 230	
TAC CAC CCT CCC AGC GCC TAC TAT GCT GTG TCC CAG CTC CGT CTG CAG	1494
Tyr His Pro Pro Ser Ala Tyr Tyr Ala Val Ser Gln Leu Arg Leu Gln	
235 240 245	
GGG AGC TGC TTC TGT CAC GGC CAT GCT GAT CGC TGC GCA CCC AAG CCT	1542
Gly Ser Cys Phe Cys His Gly His Ala Asp Arg Cys Ala Pro Lys Pro	
250 255 260	
GGG GCC TCT GCA GGC CCC TCC ACC GCT GTG CAG GTC CAC GAT GTC TGT	1590
Gly Ala Ser Ala Gly Pro Ser Thr Ala Val Gln Val His Asp Val Cys	
265 270 275	
GTC TGC CAG CAC AAC ACT GCC GGC CCA AAT TGT GAG CGC TGT GCA CCC	1638
Val Gln Cys Glu Ser Thr Ala Cys Pro Asn Cys Cys Ala Arg Pro	
TTC TAC AAC AAC UGG CCC TGG AGA UGG GGG GAG GGC CAG GAC GGC CAT	1686
Phe Tyr Asn Asn Arg Pro Trp Arg Pro Ala Glu Gly Gln Asp Ala His	

300	305	310	
GAA TGC CAA AGG TGC GAC TGC AAT GGG CAC TCA GAG ACA TGT CAC TTT			1734
Glu Cys Gln Arg Cys Asp Cys Asn Gly His Ser Glu Thr Cys His Phe			
315	320	325	
GAC CCC GCT GTG TTT GCC GCC AGC CAG GGG GCA TAT GGA GGT GTG TGT			1782
Asp Pro Ala Val Phe Ala Ala Ser Gln Gly Ala Tyr Gly Gly Val Cys			
330	335	340	
GAC AAT TGC CGG GAC CAC ACC GAA GGC AAG AAC TGT GAG CGG TGT CAG			1830
Asp Asn Cys Arg Asp His Thr Glu Gly Lys Asn Cys Glu Arg Cys Gln			
345	350	355	
CTG CAC TAT TTC CGG AAC CGG CGC CCG GGA GCT TCC ATT CAG GAG ACC			1878
Leu His Tyr Phe Arg Asn Arg Arg Pro Gly Ala Ser Ile Gln Glu Thr			
360	365	370	375
TGC ATC TCC TGC GAG TGT GAT CCG GAT GGG GCA GTG CCA GGG GCT CCC			1926
Cys Ile Ser Cys Glu Cys Asp Pro Asp Gly Ala Val Pro Gly Ala Pro			
380	385	390	
TGT GAC CCA GTG ACC GGG CAG TGT GTG TGC AAG GAG CAT GTG CAG GGA			1974
Cys Asp Pro Val Thr Gly Gln Cys Val Cys Lys Glu His Val Gln Gly			
395	400	405	
GAG CGC TGT GAC CTA TGC AAG CCG GGC TTC ACT GGA CTC ACC TAC GCC			2022
Glu Arg Cys Asp Leu Cys Lys Pro Gly Phe Thr Gly Leu Thr Tyr Ala			
410	415	420	
AAC CCG CAG GGC TGC CAC CGC TGT GAC TGC AAC ATC CTG GGG TCC CGG			2070
Asn Pro Gln Gly Cys His Arg Cys Asp Cys Asn Ile Leu Gly Ser Arg			
425	430	435	
AGG GAC ATG CCG TGT GAC GAG GAG AGT GGG CGC TGC CTT TGT CTG CCC			2118
Arg Asp Met Pro Cys Asp Glu Glu Ser Gly Arg Cys Leu Cys Leu Pro			
440	445	450	455
AAC GTG GTG GGT CCC AAA TGT GAC CAG TGT GCT CCC TAC CAC TGG AAG			2166

Asn	Val	Val	Gly	Pro	Lys	Cys	Asp	Gln	Cys	Ala	Pro	Tyr	His	Trp	Lys	
460							465							470		
CTG	GCC	AGT	GGC	CAG	GGC	TGT	GAA	CCG	TGT	GCC	TGC	GAC	CCG	CAC	AAC	2214
Leu	Ala	Ser	Gly	Gln	Gly	Cys	Glu	Pro	Cys	Ala	Cys	Asp	Pro	His	Asn	
475							480							485		
TCC	CTC	AGC	CCA	CAG	TGC	AAC	CAG	TTC	ACA	GGG	CAG	TGC	CCC	TGT	CGG	2262
Ser	Leu	Ser	Pro	Gln	Cys	Asn	Gln	Phe	Thr	Gly	Gln	Cys	Pro	Cys	Arg	
490							495							500		
GAA	GGC	TTT	GGT	GGC	CTG	ATG	TGC	AGC	GCT	GCA	GCC	ATC	CGC	CAG	TGT	2310
Glu	Gly	Phe	Gly	Gly	Leu	Met	Cys	Ser	Ala	Ala	Ala	Ile	Arg	Gln	Cys	
505							510							515		
CCA	GAC	CGG	ACC	TAT	GGA	GAC	GTG	GCC	ACA	GGA	TGC	CGA	GCC	TGT	GAC	2358
Pro	Asp	Arg	Thr	Tyr	Gly	Asp	Val	Ala	Thr	Gly	Cys	Arg	Ala	Cys	Asp	
520							525							535		
TGT	GAT	TTC	CGG	GGA	ACA	GAG	GGC	CCG	GGC	TGC	GAC	AAG	GCA	TCA	GGC	2406
Cys	Asp	Phe	Arg	Gly	Thr	Glu	Gly	Pro	Gly	Cys	Asp	Lys	Ala	Ser	Gly	
540							545							550		
CGC	TGC	CTC	TGC	CGC	CCT	GGC	TTG	ACC	GGG	CCC	CGC	TGT	GAC	CAG	TGC	2454
Arg	Cys	Leu	Cys	Arg	Pro	Gly	Leu	Thr	Gly	Pro	Arg	Cys	Asp	Gln	Cys	
555							560							565		
CAG	CGA	GGC	TAC	TGC	AAT	CGC	TAC	CCG	GTG	TGC	GTG	GCC	TGC	CAC	CCT	2502
Gln	Arg	Gly	Tyr	Cys	Asn	Arg	Tyr	Pro	Val	Cys	Val	Ala	Cys	His	Pro	
570							575							580		
TGC	TTC	CAG	ACC	TAT	GAT	GCG	GAC	CTC	CGG	GAG	CAG	GCC	CTG	CGC	TTT	2550
Cys	Phe	Gln	Thr	Tyr	Asp	Ala	Asp	Leu	Arg	Glu	Gln	Ala	Leu	Arg	Phe	
585							590							595		
Gly	Arg	Leu	Arg	Asn	Ala	Thr	Ala	Ser	Leu	Trp	Ser	Gly	Pro	Gly	Leu	
600							605							610		
														615		

GAG GAC CGT GGC CTG GCC TCC CGG ATC CTA GAT GCA AAG AGT AAG ATT		2646	
Glu Asp Arg Gly Leu Ala Ser Arg Ile Leu Asp Ala Lys Ser Lys Ile			
620	625	630	
GAG CAG ATC CGA GCA GTT CTC AGC AGC CCC GCA GTC ACA GAG CAG GAG		2694	
Glu Gln Ile Arg Ala Val Leu Ser Ser Pro Ala Val Thr Glu Gln Glu			
635	640	645	
GTG GCT CAG GTG GCC AGT GCC ATC CTC TCC CTC AGG CGA ACT CTC CAG		2742	
Val Ala Gln Val Ala Ser Ala Ile Leu Ser Leu Arg Arg Thr Leu Gln			
650	655	660	
GGC CTG CAG CTG GAT CTG CCC CTG GAG GAG GAG ACG TTG TCC CTT CCG		2790	
Gly Leu Gln Leu Asp Leu Pro Leu Glu Glu Glu Thr Leu Ser Leu Pro			
665	670	675	
AGA GAC CTG GAG AGT CTT GAC AGA AGC TTC AAT GGT CTC CTT ACT ATG		2838	
Arg Asp Leu Glu Ser Leu Asp Arg Ser Phe Asn Gly Leu Leu Thr Met			
680	685	690	695
TAT CAG AGG AAG AGG GAG CAG TTT GAA AAA ATA AGC AGT GCT GAT CCT		2886	
Tyr Gln Arg Lys Arg Glu Gln Phe Glu Lys Ile Ser Ser Ala Asp Pro			
700	705	710	
TCA GGA GCC TTC CGG ATG CTG AGC ACA GCC TAC GAG CAG TCA GCC CAG		2934	
Ser Gly Ala Phe Arg Met Leu Ser Thr Ala Tyr Glu Gln Ser Ala Gln			
715	720	725	
GCT GCT CAG CAG GTC TCC GAC AGC TCG CGC CTT TTG GAC CAG CTC AGG		2982	
Ala Ala Gln Gln Val Ser Asp Ser Ser Arg Leu Leu Asp Gln Leu Arg			
730	735	740	
GAC AGC CGG AGA GAG GCA GAG AGG CTG GTG CGG CAG GCG GGA GGA GGA		3030	
Asp Ser Arg Arg Glu Ala Glu Arg Leu Val Arg Gln Ala Gly Gly Gly			
745	750	755	
GGA GGC ACC GGC AGC CCC AAG CTT GTG GCC CTG AGG CTG GAG ATG TCT		3078	
Gly Gly Thr Gly Ser Pro Lys Leu Val Ala Leu Arg Leu Glu Met Ser			

特平 8-243060

760	765	770	775
TCG TTG CCT GAC CTG ACA CCC ACC TTC AAC AAG CTC TGT GGC AAC TCC			3126
Ser Leu Pro Asp Leu Thr Pro Thr Phe Asn Lys Leu Cys Gly Asn Ser			
780	785	790	
AGG CAG ATG GCT TGC ACC CCA ATA TCA TGC CCT GGT GAG CTA TGT CCC			3174
Arg Gln Met Ala Cys Thr Pro Ile Ser Cys Pro Gly Glu Leu Cys Pro			
795	800	805	
CAA GAC AAT GGC ACA GCC TGT GGC TCC CGC TGC AGG GGT GTC CTT CCC			3222
Gln Asp Asn Gly Thr Ala Cys Gly Ser Arg Cys Arg Gly Val Leu Pro			
810	815	820	
AGG GCC GGT GGG GCC TTC TTG ATG GCG GGG CAG GTG GCT GAG CAG CTG			3270
Arg Ala Gly Gly Ala Phe Leu Met Ala Gly Gln Val Ala Glu Gln Leu			
825	830	835	
CGG GGC TTC AAT GCC CAG CTC CAG CGG ACC AGG CAG ATG ATT AGG GCA			3318
Arg Gly Phe Asn Ala Gln Leu Gln Arg Thr Arg Gln Met Ile Arg Ala			
840	845	850	855
GCC GAG GAA TCT GCC TCA CAG ATT CAA TCC AGT GCC CAG CGC TTG GAG			3366
Ala Glu Glu Ser Ala Ser Gln Ile Gln Ser Ser Ala Gln Arg Leu Glu			
860	865	870	
ACC CAG GTG AGC GCC AGC CGC TCC CAG ATG GAG GAA GAT GTC AGA CGC			3414
Thr Gln Val Ser Ala Ser Arg Ser Gln Met Glu Glu Asp Val Arg Arg			
875	880	885	
ACA CGG CTC CTA ATC CAG CAG GTC CGG GAC TTC CTA ACA GAC CCC GAC			3462
Thr Arg Leu Leu Ile Gln Gln Val Arg Asp Phe Leu Thr Asp Pro Asp			
890	895	900	
905	910	915	
TGG CTG CCC ACA GAC TCA GCT ACT GTT CTG CAG AAG ATG AAT GAG ATC			3558

Trp Leu Pro Thr Asp Ser Ala Thr Val Leu Gln Lys Met Asn Glu Ile
 920 925 930 935
 CAG GCC ATT GCA GCC AGG CTC CCC AAC GTG GAC TTG GTG CTG TCC CAG 3606
 Gln Ala Ile Ala Ala Arg Leu Pro Asn Val Asp Leu Val Leu Ser Gln
 940 945 950
 ACC AAG CAG GAC ATT GCG CGT GCC CGC CGG TTG CAG GCT GAG GCT GAG 3654
 Thr Lys Gln Asp Ile Ala Arg Ala Arg Arg Leu Gln Ala Glu Ala Glu
 955 960 965
 GAA GCC AGG AGC CGA GCC CAT GCA GTG GAG GGC CAG GTG GAA GAT GTG 3702
 Glu Ala Arg Ser Arg Ala His Ala Val Glu Gly Gln Val Glu Asp Val
 970 975 980
 GTT GGG AAC CTG CGG CAG GGG ACA GTG GCA CTG CAG GAA GCT CAG GAC 3750
 Val Gly Asn Leu Arg Gln Gly Thr Val Ala Leu Gln Glu Ala Gln Asp
 985 990 995
 ACC ATG CAA GGC ACC AGC CGC TCC CTT CGG CTT ATC CAG GAC AGG GTT 3798
 Thr Met Gln Gly Thr Ser Arg Ser Leu Arg Leu Ile Gln Asp Arg Val
 1000 1005 1010 1015
 GCT GAG GTT CAG CAG GTA CTG CGG CCA GCA GAA AAG CTG GTG ACA AGC 3846
 Ala Glu Val Gln Gln Val Leu Arg Pro Ala Glu Lys Leu Val Thr Ser
 1020 1025 1030
 ATG ACC AAG CAG CTG GGT GAC TTC TGG ACA CGG ATG GAG GAG CTC CGC 3894
 Met Thr Lys Gln Leu Gly Asp Phe Trp Thr Arg Met Glu Glu Leu Arg
 1035 1040 1045
 CAC CAA GCC CGG CAG CAG GGG GCA GAG GCA GTC CAG GCC CAG CTT 3942
 His Gln Ala Arg Gln Gln Gly Ala Glu Ala Val Gln Ala Gln Gln Leu
 1050 1055 1060
 GCG GAA GGT GCC AGC GAG CAG GCA TTG AGT GCC CAA GAG GGA TTT GAG 3990
 Ala Glu Gly Ala Ser Glu Gln Ala Leu Ser Ala Gln Glu Gly Phe Glu
 1065 1070 1075

AGA ATA AAA CAA AAG TAT GCT GAG TTG AAG GAC CGG TTG GGT CAG AGT	4038		
Arg Ile Lys Gln Lys Tyr Ala Glu Leu Lys Asp Arg Leu Gly Gln Ser			
1080	1085	1090	1095
TCC ATG CTG GGT GAG CAG GGT GCC CGG ATC CAG AGT GTG AAG ACA GAG	4086		
Ser Met Leu Gly Glu Gln Gly Ala Arg Ile Gln Ser Val Lys Thr Glu			
1100	1105	1110	
GCA GAG GAG CTG TTT CGG GAG ACC ATG GAG ATG ATG GAC AGG ATG AAA	4134		
Ala Glu Glu Leu Phe Gly Glu Thr Met Glu Met Met Asp Arg Met Lys			
1115	1120	1125	
GAC ATG GAG TTG GAG CTG CTG CGC CCC AGC CAG GCC ATC ATG CTG CGC	4182		
Asp Met Glu Leu Glu Leu Leu Arg Gly Ser Gln Ala Ile Met Leu Arg			
1130	1135	1140	
TCA GCG GAC CTG ACA GGA CTG GAG AAG CGT GTG GAG CAG ATC CGT GAC	4230		
Ser Ala Asp Leu Thr Gly Leu Glu Lys Arg Val Glu Gln Ile Arg Asp			
1145	1150	1155	
CAC ATC AAT GGG CGC GTG CTC TAC TAT GCC ACC TGC AAG T	4270		
His Ile Asn Gly Arg Val Leu Tyr Tyr Ala Thr Cys Lys			
1160	1165	1170	
GATGCTACAG CTTCCAGCCC GTTGCCAC TCATCTGCCG CCTTGCTT TGGTTGGGG	4330		
CAGATTGGGT TGGAAATGCTT TCCATCTCCA GGAGACTTTC ATGCAGCCTA AAGTACAGCC	4390		
TGGACCACCC CTGGTGTGTA GCTAGTAAGA TTACCCCTGAG CTGCAGCTGA GCCTGAGCCA	4450		
ATGGGACAGT TACACTTGAC AGACAAAGAT GGTGGAGATT GGCATGCCAT TGAAACTAAG	4510		
AGCTCTCAAG TCAAGGAAGC TGGGCTGGGC AGTATCCCC GCCTTAGTT CTCCACTGGG	4570		
GAGGAATCCT GGACCAAGCA CAAAAACTTA ACAAAAGTGA TGTAAAAATG AAAAGCCAAA	4630		
TAAAAATCTT TGGAAAAGAG CCTGGAGGTT CAACGAG	4667		

10000

配列番号：26

配列の長さ：1086

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：HP10298

配列の特徴：

特徴を表す記号：CDS

存在位置：138..506

特徴を決定した方法：E

配列

TTAATTCC CCGAAATCAG ACTGCTGCCT TGGACCGGGA CAGCTCGCGG CCCCCGAGAG	60	
CTCTAGCCGT CGAGGAGCTG CCTGGGGACG TTTGCCCTGG GGCCCCAGCC TGGCCCGGGT	120	
CACCCTGGCA TGAGGAG ATG GGC CTG TTG CTC CTG GTC CCA TTG CTC CTG	170	
Met Gly Leu Leu Leu Leu Val Pro Leu Leu Leu		
1	5	10
CTG CCC GGC TCC TAC GGA CTG CCC TTC TAC AAC GGC TTC TAC TAC TCC	218	
Leu Pro Gly Ser Tyr Gly Leu Pro Phe Tyr Asn Gly Phe Tyr Tyr Ser		
15	20	25
AAC AGC GCC AAC GAC CAG AAC CTA GGC AAC GGT CAT GGC AAA GAC CTC	266	
Asn Ser Ala Asn Asp Gln Asn Leu Gly Asn Gly His Gly Lys Asp Leu		
30	35	40
CTT AAT GGA GTG AAG CTG GTG GAG ACA CCC GAG GAG ACC CTG TTC	314	
Leu Asn Gly Val Lys Leu Val Val Glu Thr Pro Glu Glu Thr Leu Phe		
45	50	55

ACC CGC ATC CTA ACT GTG GGC CCC CAG AGC CTG GGG TCC GAA GCT TTG	362
Thr Arg Ile Leu Thr Val Gly Pro Gln Ser Leu Gly Ser Glu Ala Leu	
60 65 70 75	
GCT TCC CCG ACC CGC AGA GCC GCT TGT ACG GTG TTT ACT GCT ACC GCC	410
Ala Ser Pro Thr Arg Arg Ala Ala Cys Thr Val Phe Thr Ala Thr Ala	
80 85 90	
AGC ACT AGG ACC TGG GGC CCT CCC CTG CCG CAT TCC CTC ACT GGC TGT	458
Ser Thr Arg Thr Trp Gly Pro Pro Leu Pro His Ser Leu Thr Gly Cys	
95 100 105	
GTA TTT ATT GAG TGG TTC GTT TTC CCT TGT CGG TTG GAG CCA TTT	503
Val Phe Ile Glu Trp Phe Val Phe Pro Cys Gly Leu Glu Pro Phe	
110 115 120	
TAACGT TTTTATACTT CTCAATTAA ATTTCCTTA AACATTTTT TACTATTTT	560
TGTAAAGCAA ACAGAACCCA ATGCCTCCCT TTGCTCCTGG ATGCCCACT CCAGGAATCA	620
TGCTTGCTCC CCTGGGCCAT TTGCGGTTTT GTGGGCTTCT GGAGGGTTCC CCGCCATCCA	680
GGCTGGTCTC CCTCCCTTAA GGAGGTTGGT GCCCAGAGTG GGCGGTGGCC TGTCTAGAAT	740
GCCGCCGGGA GTCCGGGCAT GGTGGGCACA GTTCTCCCTG CCCCTCAGCC TGGGGAAAGA	800
AGAGGGCCTC GGGGGCCTCC GGAGCTGGC TTTGGGCCTC TCCTGCCAC CTCTACTTCT	860
CTGTGAAGCC GCTGACCCCA GTCTGCCAC TGAGGGCTA GGGCTGGAAG CCAGTTCTAG	920
GCTTCCAGGC GAAAGCTGAG GGAAGGAAGA AACTCCCCTC CCCGTTCCCC TTCCCTCTC	980
GGTTCCAAAG AATCTGTTT GTTGTCAATT GTTCTCCTG TTTCCCTGTG TGGGGAGGGG	1040
CCCTCAGGTG TGTGTACTTT GGACAATAAA TGGTGCTATG ACTGCC	1086

【0097】

特平 8-243060

配列番号：27

配列の長さ： 866

配列の型：核酸

鎖の数： 二本鎖

トポロジー：直鎖状

配列の種類： cDNA to mRNA

起源：

生物名：ホモ＝サピエンス

細胞の種類：胃癌

クローン名：HP10368

配列の特徴：

特徴を表す記号： CDS

存在位置： 73.. 600

特徴を決定した方法： E

配列

ACTCAGAAC	TTGGACCGCA	TCCTAGCCGC	CGACTCACAC	AAGGCAGGTG	GGTGAGGAAA	60									
TCCAGAGTTG	CC	ATG	GAG	AAA	ATT	CCA	GTG	TCA	GCA	TTC	TTG	CTC	CTT	GTG	111
Met Glu Lys Ile Pro Val Ser Ala Phe Leu Leu Leu Val															

1	5	10
---	---	----

GCC	CTC	TCC	TAC	ACT	CTG	GCC	AGA	GAT	ACC	ACA	GTC	AAA	CCT	GGA	GCC	159
Ala	Leu	Ser	Tyr	Thr	Leu	Ala	Arg	Asp	Thr	Thr	Val	Lys	Pro	Gly	Ala	
15	20	25														

AAA	AAG	GAC	ACA	AAG	GAC	TCT	CGA	CCC	AAA	CTG	CCC	CAG	ACC	CTC	TCC	207
Lys	Lys	Asp	Thr	Lys	Asp	Ser	Arg	Pro	Lys	Leu	Pro	Gln	Thr	Leu	Ser	
30	35	40	45													

AGA	GGT	TGG	GGT	GAC	CAA	CTC	ATC	TGG	ACT	CAG	ACA	TAT	GAA	GCT	255
Arg	Cys	Trp	Cys	Glu	Cys	Leu	Leu	Trp	Thr	Cys	Thr	Trp	Cys	Cys	Ala

CTA	TAT	AAA	TCC	AAG	ACA	AGC	AAU	AAA	CCC	TIG	AIG	AII	AII	CAT	CAU	303
Leu	Tyr	Lys	Ser	Lys	Thr	Ser	Asn	Lys	Pro	Leu	Met	Ile	Ile	His	His	

65	70	75	
TTG GAT GAG TGC CCA CAC AGT CAA GCT TTA AAG AAA GTG TTT GCT GAA			351
Leu Asp Glu Cys Pro His Ser Gln Ala Leu Lys Lys Val Phe Ala Glu			
80	85	90	
AAT AAA GAA ATC CAG AAA TTG GCA GAG CAG TTT GTC CTC CTC AAT CTG			399
Asn Lys Glu Ile Gln Lys Leu Ala Glu Gln Phe Val Leu Leu Asn Leu			
95	100	105	
GTT TAT GAA ACA ACT GAC AAA CAC CTT TCT CCT GAT GGC CAG TAT GTC			447
Val Tyr Glu Thr Thr Asp Lys His Leu Ser Pro Asp Gly Gln Tyr Val			
110	115	120	125
CCC AGG ATT ATG TTT GTT GAC CCA TCT CTG ACA GTT AGA GCC GAT ATC			495
Pro Arg Ile Met Phe Val Asp Pro Ser Leu Thr Val Arg Ala Asp Ile			
130	135	140	
ACT GGA AGA TAT TCA AAC CGT CTC TAT GCT TAC GAA CCT GCA GAT ACA			543
Thr Gly Arg Tyr Ser Asn Arg Leu Tyr Ala Tyr Glu Pro Ala Asp Thr			
145	150	155	
GCT CTG TTG CTT GAC AAC ATG AAG AAA GCT CTC AAG TTG CTG AAG ACT			591
Ala Leu Leu Leu Asp Asn Met Lys Lys Ala Leu Lys Leu Leu Lys Thr			
160	165	170	
GAA TTG TAAAGAAAAA AAATCTCCAA GCCCTTCTGT CTGTCAGGCC TTG			640
Glu Leu			
175			
AGACTTGAAA CCAGAAGAAG TGTGAGAAGA CTGGCTAGTG TGGAAGCATA GTGAACACAC			700
TGATTAGGTT ATGGTTAACG GTTACAACAA CTATTTTTA AGAAAAACAA GTTTAGAAA			760
TTGGTTCA AGTGTACATG TGTGAAACA ATATTGTATA CTACCATAGT GAGCCATGAT			820
TTCTAAAAAA AAAAAATAAA TGTTTGGGG GTGTTCTGTT TTCTCC			866

【0098】

【図面の簡単な説明】

【図1】 分泌シグナル配列検出ベクター pSSD3 の構造を表す図である。

【図2】 分泌シグナル配列－ウロキナーゼ融合遺伝子の作製法を示す図である。

【図3】 クローンHP00658がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図4】 クローンHP00714がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図5】 クローンHP00876がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図6】 クローンHP01134がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図7】 クローンHP10029がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図8】 クローンHP10189がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

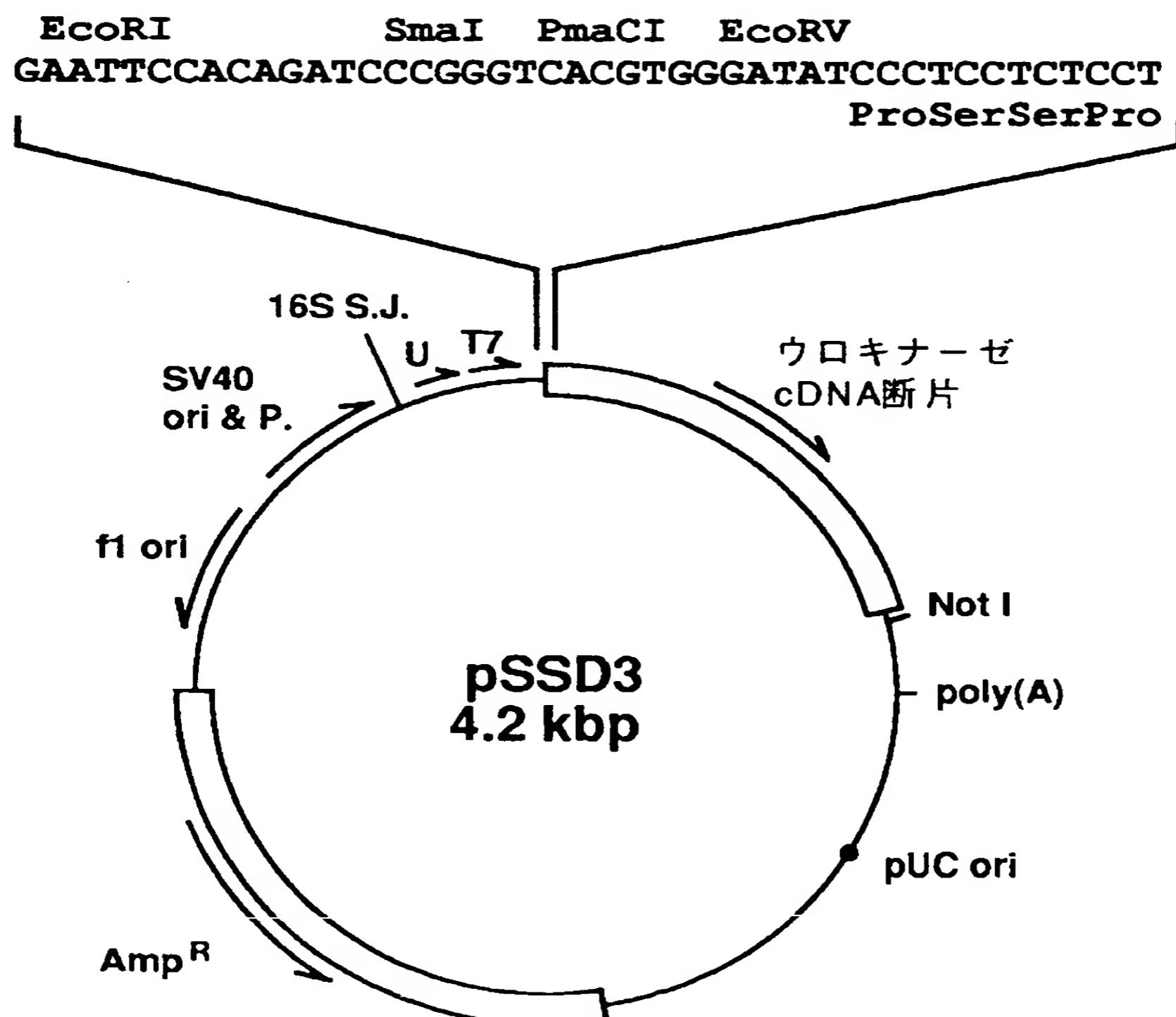
【図9】 クローンHP10269がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【図10】 クローンHP10298がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

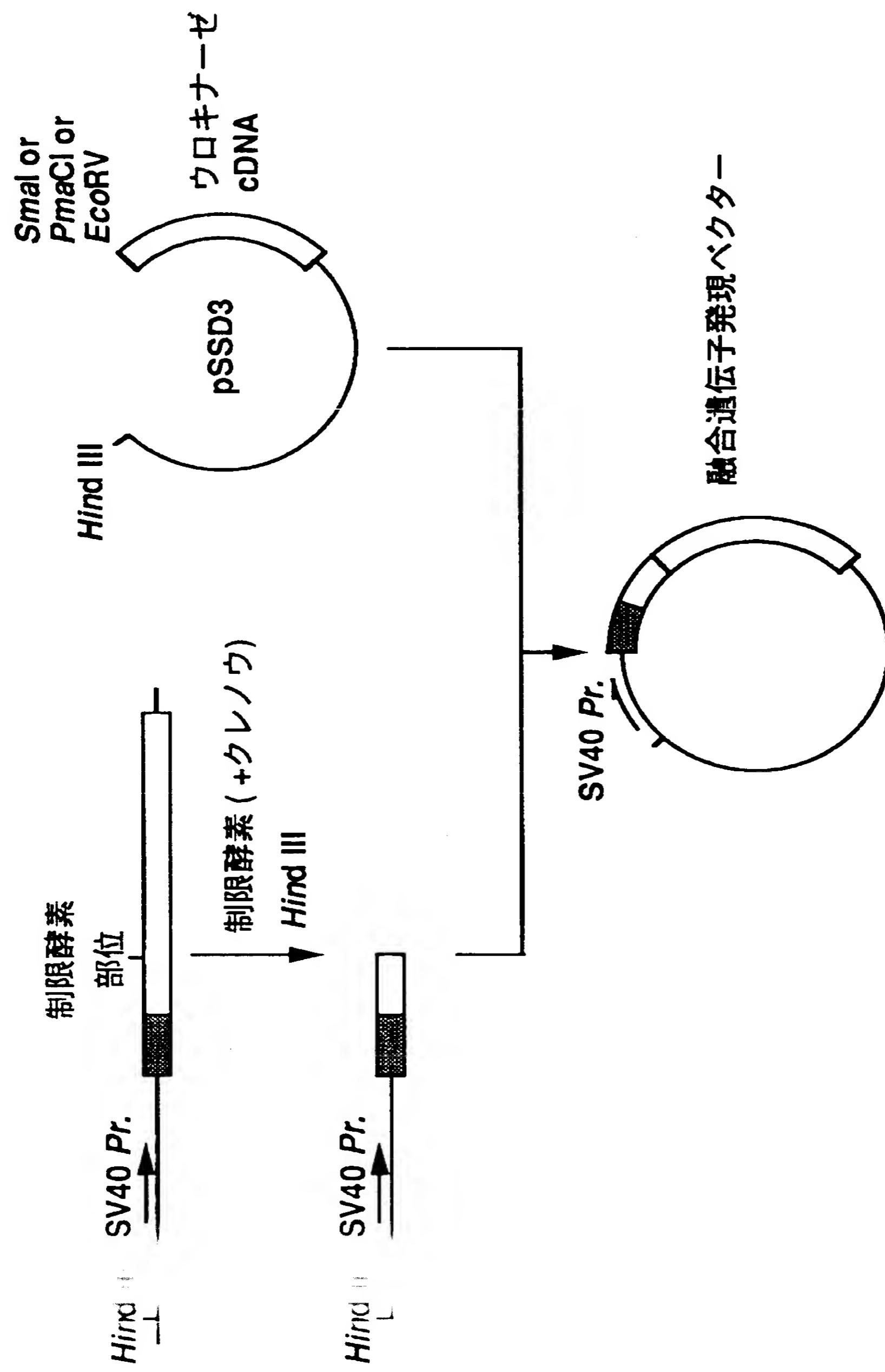
【図11】 クローンHP10368がコードする蛋白質の疎水性／親水性プロファイルを示す図である。

【書類名】 図面

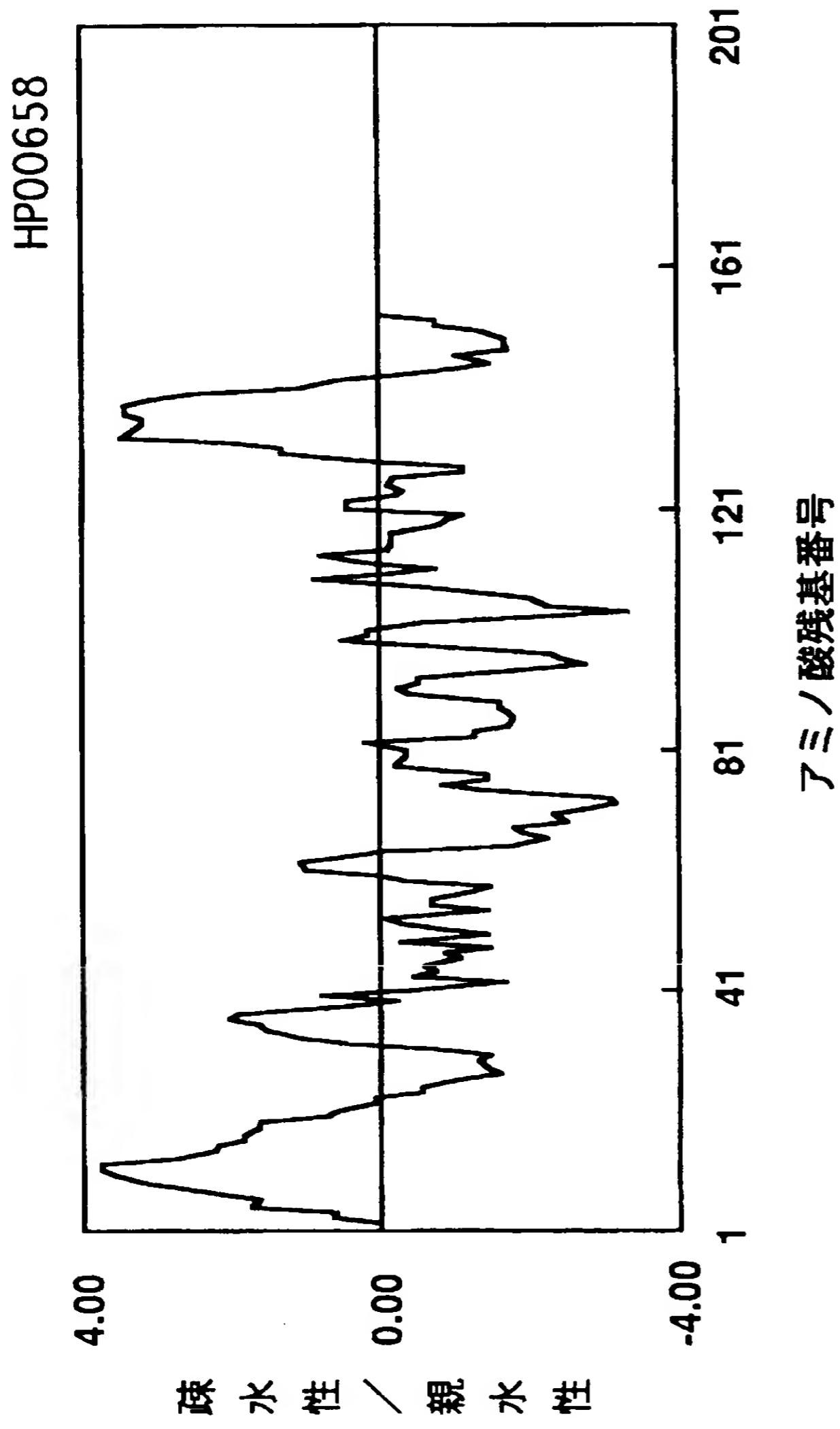
【図1】



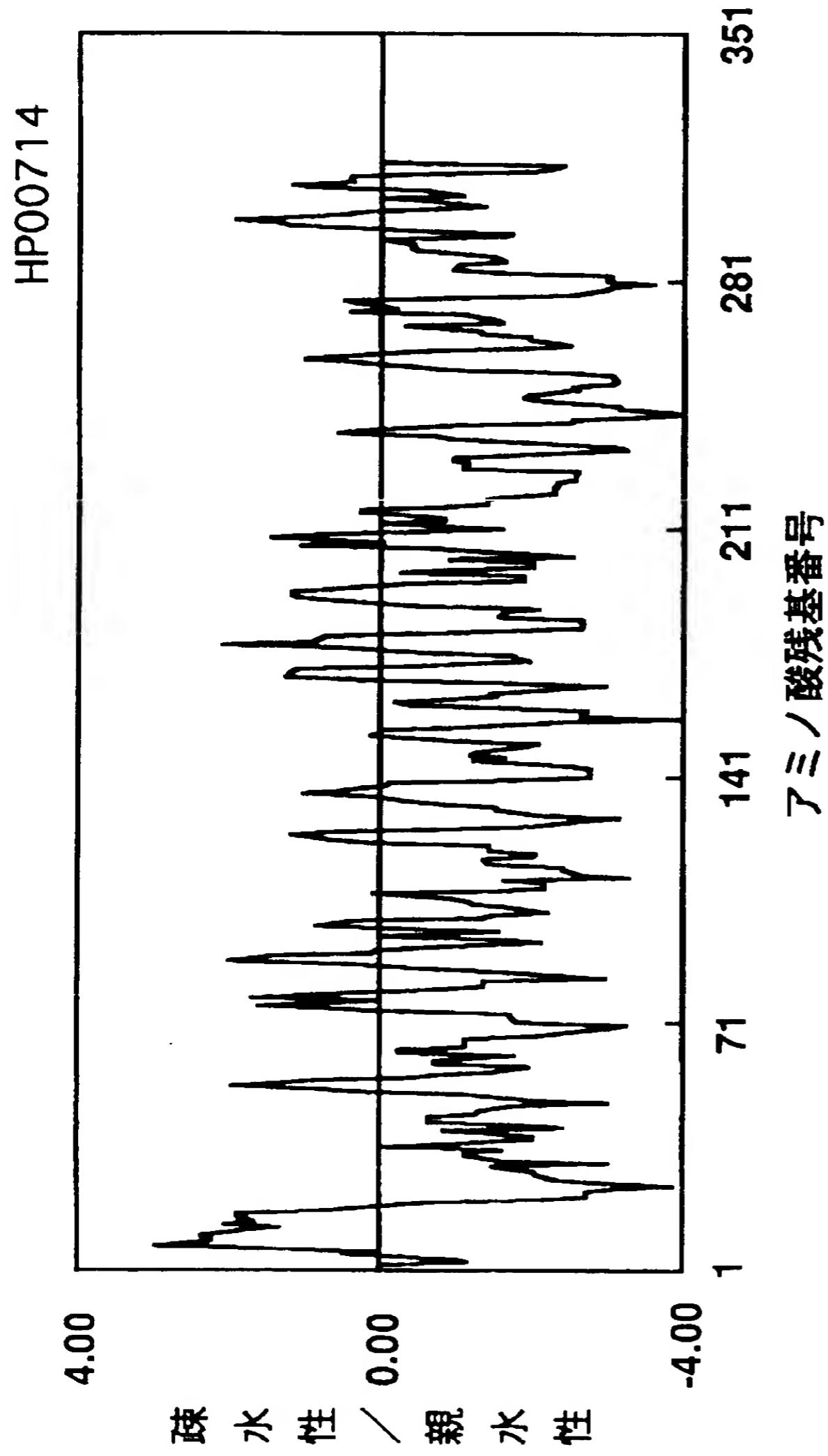
【図2】



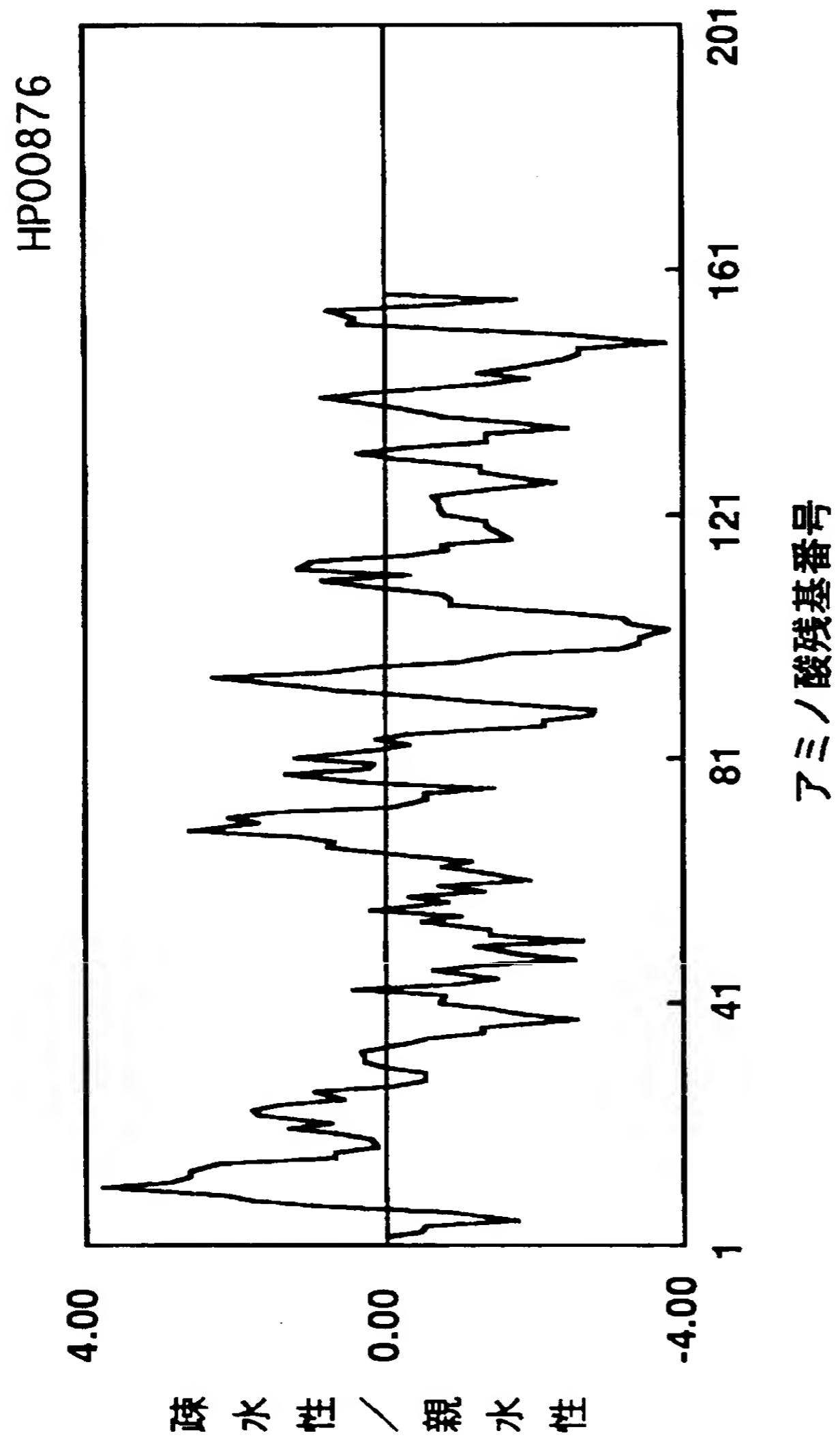
〔図3〕



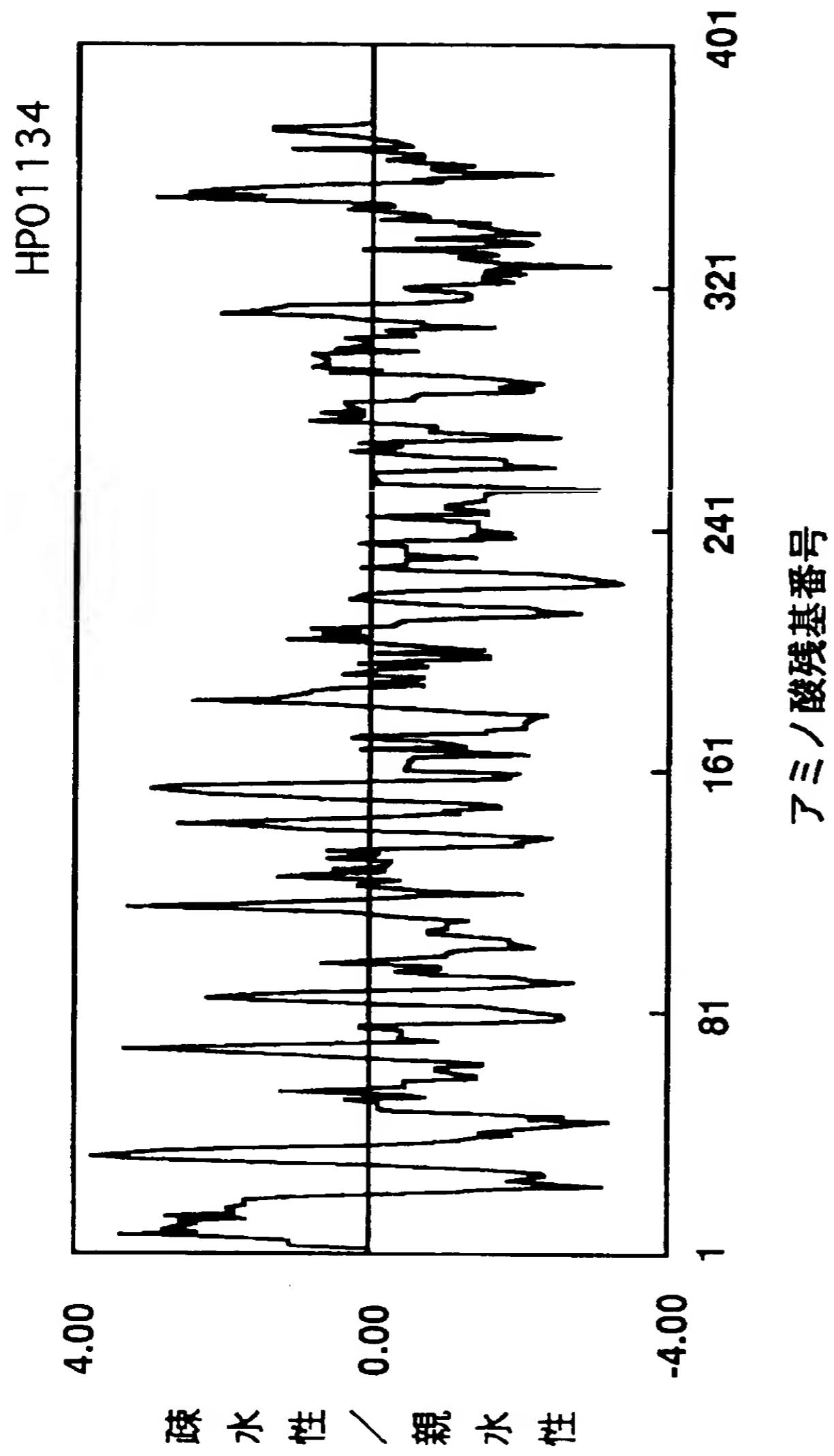
【図4】



【図5】

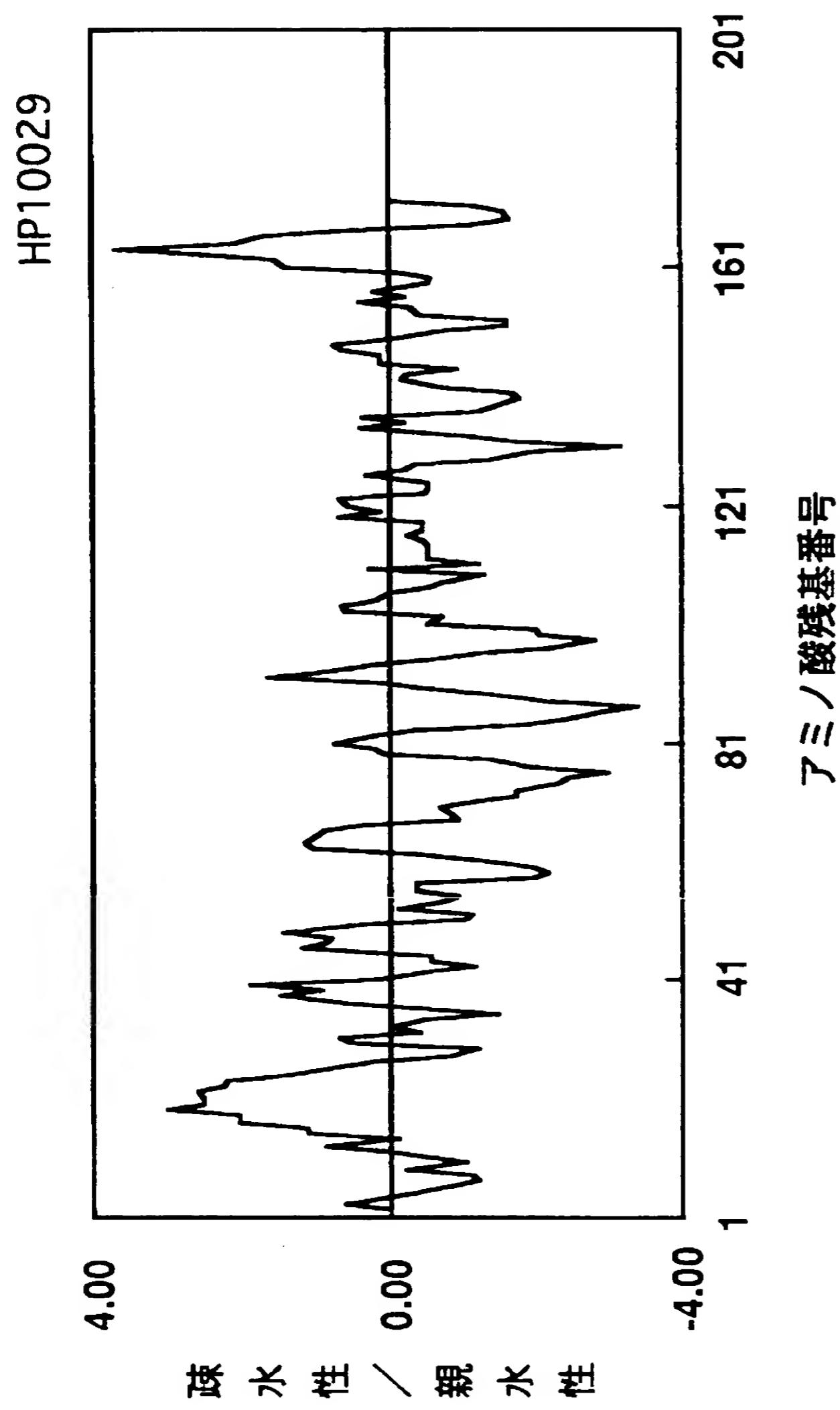


【図6】

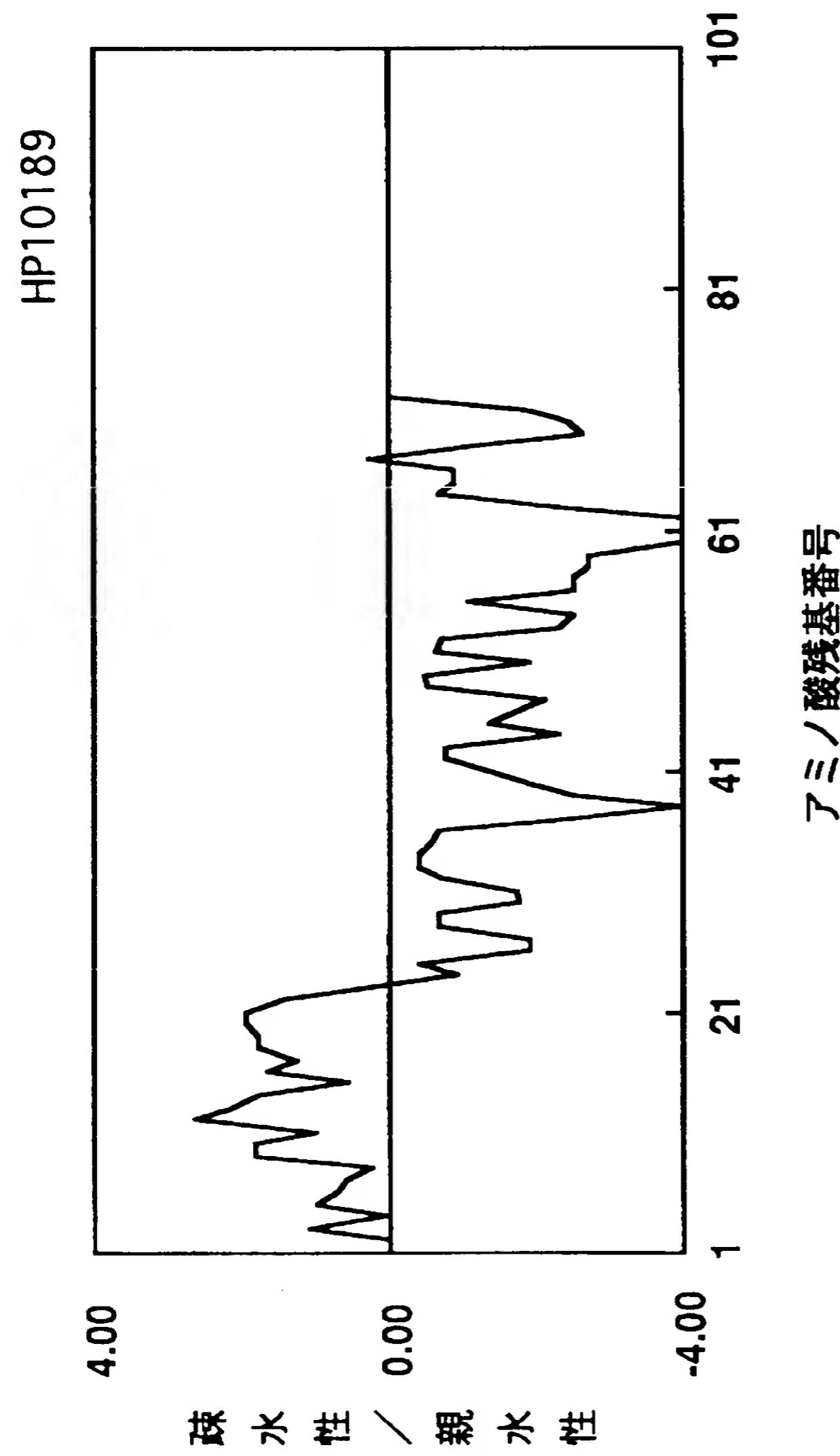


特平 8-243060

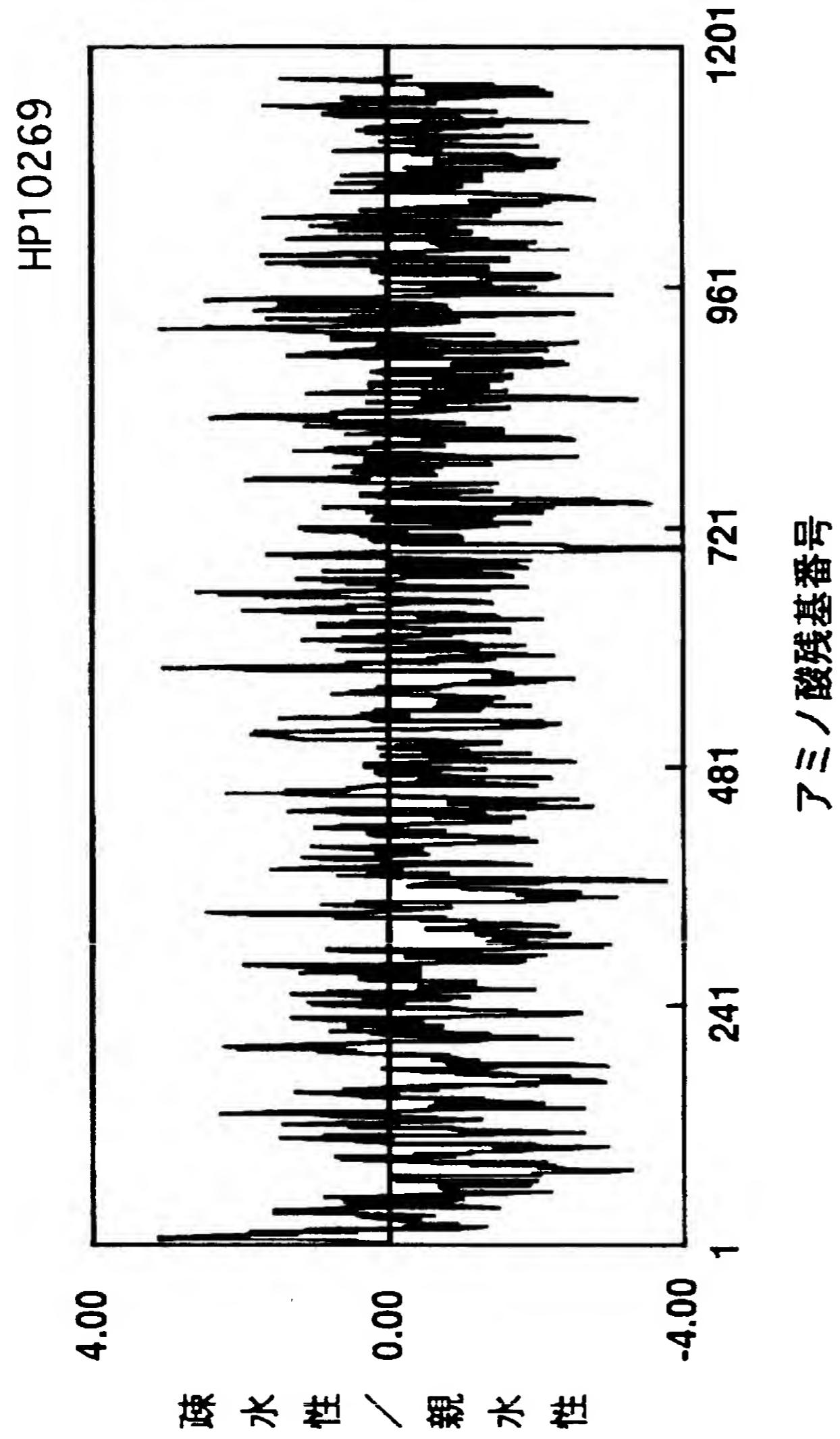
【図7】



【図8】

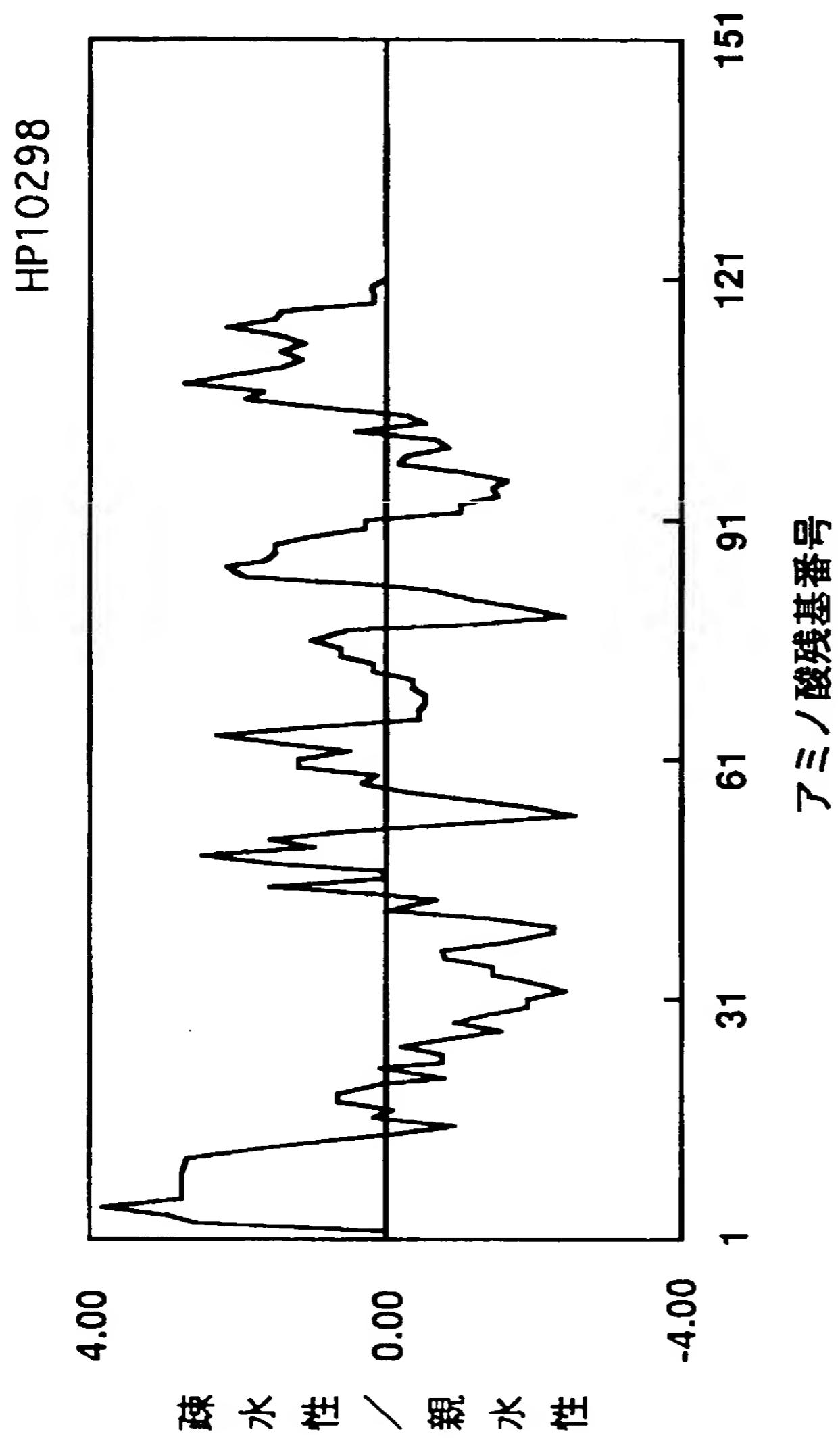


【図9】



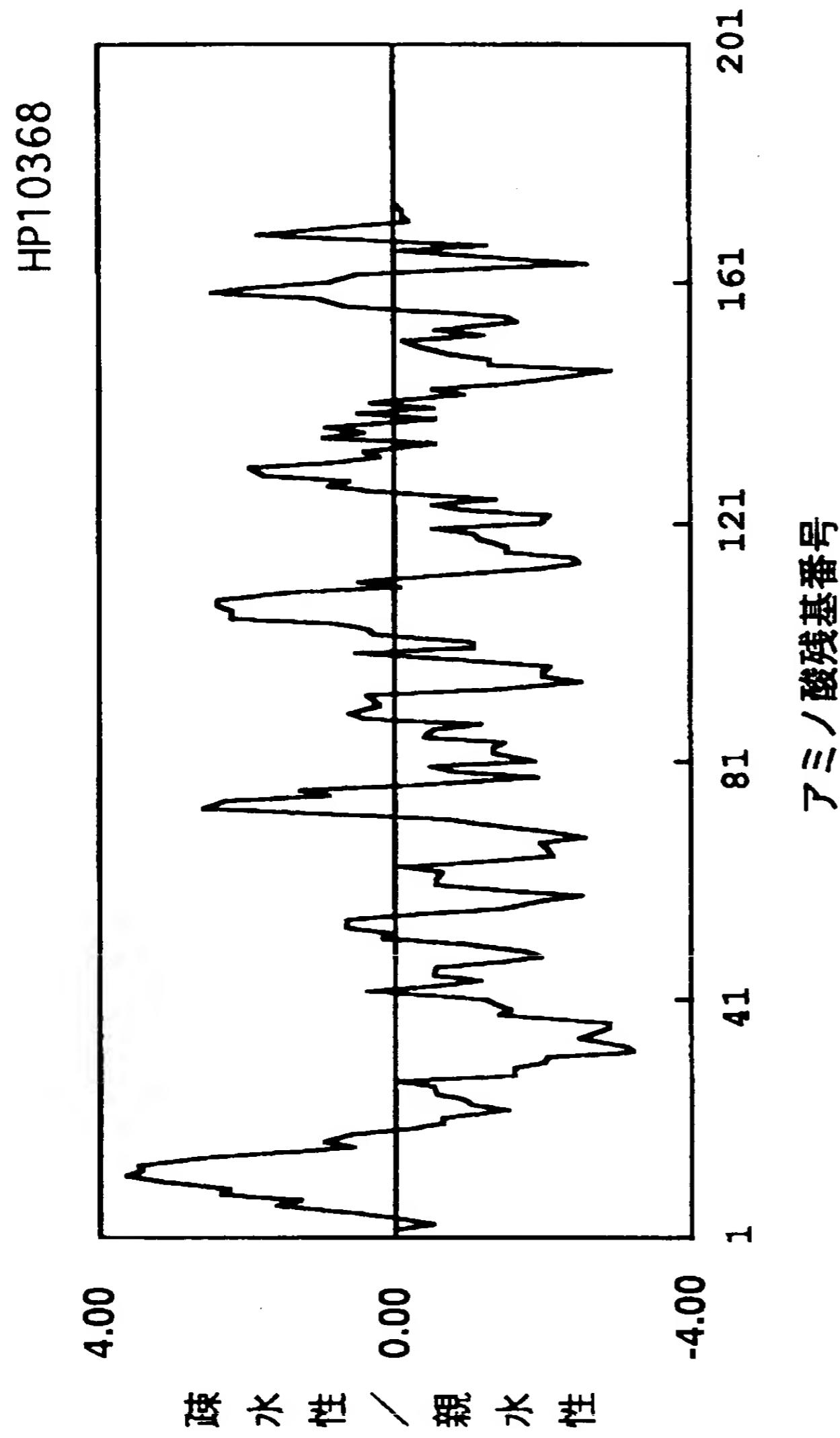
特平 8-243060

【図10】



特平 8-243060

【図 1 1】



【書類名】 要約書

【要約】

【課題】 分泌シグナル配列を有するヒト蛋白質、およびそれをコードしている c DNA を提供する。

【解決手段】 配列番号 1 から配列番号 9 で表されるアミノ酸配列を含む蛋白質および該蛋白質をコードする DNA、例えば配列番号 10 から配列番号 18 で表される塩基配列を含む c DNA。分泌機能が確認された分泌シグナル配列を有するヒト蛋白質をコードしている c DNA、およびこのヒト c DNA の組換え体を発現させることにより該蛋白質を提供することができる。

【選択図】 なし

【書類名】 職権訂正データ
【訂正書類】 特許願

＜認定情報・付加情報＞

【特許出願人】 申請人 000173762
【識別番号】
【住所又は居所】 神奈川県相模原市西大沼4丁目4番1号
【氏名又は名称】 財団法人相模中央化学研究所
【特許出願人】 596134998
【識別番号】
【住所又は居所】 東京都目黒区中町2丁目20番3号
【氏名又は名称】 株式会社プロテジーン

出願人履歴情報

識別番号 [000173762]

1. 変更年月日 1995年 4月14日
[変更理由] 住所変更
住 所 神奈川県相模原市西大沼4丁目4番1号
氏 名 財団法人相模中央化学研究所

出願人履歴情報

識別番号 [596134998]

1. 変更年月日 1996年 9月13日
[変更理由] 新規登録
住 所 東京都目黒区中町2丁目20番3号
氏 名 株式会社プロテジーン